

Communication

Information médias théories pratiques

Vol. 42/2 | 2025

Vol. 42/2

La datafication de la société : un état des lieux des études critiques sur les données

Florence Millerand, Mélanie Millette et Lucie Delias (dir.)



Édition électronique

URL : <https://journals.openedition.org/communication/20957>

ISSN : 1920-7344

Éditeur

Université Laval

Référence électronique

Florence Millerand, Mélanie Millette et Lucie Delias (dir.), *Communication*, Vol. 42/2 | 2025, « Vol. 42/2 » [En ligne], mis en ligne le 01 décembre 2025, consulté le 11 décembre 2025. URL : <https://journals.openedition.org/communication/20957>

Ce document a été généré automatiquement le 11 décembre 2025.



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont susceptibles d'être soumis à des autorisations d'usage spécifiques.

SOMMAIRE

Dossier

La datafication de la société

Un état des lieux des études critiques sur les données
Florence Millerand, Lucie Delias et Mélanie Millette

Articles

La gouvernance de la « civilisation écologique » en Chine

Vers une datafication de l'intégrité écologique des entreprises et des individus
Élodie René et Samuel Tronçon

Pratiques numériques et mise en données de la défense des droits humains

Vers une datafication de nos affects ?
Normand Landry, Julien Hocine, Simon Fitzbay et Marie-Josée Dupuis

Vie de laboratoire urbain et quartiers intelligents

Analyse des trajectoires de Sidewalk Toronto et du Quartier des spectacles de Montréal
Etienne Grenier et Nicolas Chartier-Edwards

Pour l'amour de la donnée

Vers une datafication de nos affects ?
Camille Alloing, Ambre Marionneau, Julien Pierre et Fabien Richert

Promesses technomnésiques et datafication de la mémoire personnelle

Un état des lieux des études critiques sur les données
Rémi Rouge, Emmanuelle Caccamo et Katharina Niemeyer

La datafication des soins en maisons de santé

Un pouvoir professionnel inégalement distribué
Manon Plegat

Datafication d'un corps faillible et pratiques d'autosoins

Explorer les data valences pour penser la sur-veillance
Sylvie Grosjean, Fabienne Martin-Juchât et Ambre Davat

Donnée administrative, ouvre-toi !

Le travail scientifique sur des données de subventions associatives
Cécile Rodrigues et Thomas Chevallier

Cultures de la quantification et quantification de la culture

Les effets de la mise en nombre des pratiques littéraires sur les utilisateurs de plateformes numériques
Juliette Parmentier

Lutter contre les violences des données issues de la datafication

Pour une littératie des données queer
Ugo Verdi

L'artisanat des données à l'ère de la datafication

François Lambotte et Kevin Carillon

Les études critiques sur les données

Généalogie et cartographie d'un champ
Florence Millerand, Mélanie Millette et Lucie Delias

Lectures

Yves WINKIN et Jean-Marie CHARPENTIER (2025), La communication au long cours.

Conversations sur les sciences de la communication

Caen, Éditions C&F, Coll. « Recherche »

Alexandre Eyries

Dossier

La datafication de la société

Un état des lieux des études critiques sur les données

Florence Millerand, Lucie Delias et Mélanie Millette

- 1 Les *data* ont, plus que jamais, le vent en poupe. Depuis la vague des données massives (ou *big data*) jusqu'aux récents développements en matière d'intelligence artificielle (IA) générative, qui ont donné lieu aux robots conversationnels du type ChatGPT, les données sont désormais considérées comme une ressource stratégique majeure pour le devenir de nos sociétés contemporaines toujours plus automatisées. Dans ce contexte, il faut rappeler que le fonctionnement des systèmes basés sur l'IA, tout comme les grandes plateformes de médias sociaux et les algorithmes sur lesquels ils reposent, repose sur l'analyse automatisée de masses gigantesques données. Ces données sont de nature et de provenance variées. Elles sont produites par le biais de diverses activités, y compris par les individus dans leur vie quotidienne (consulter les nouvelles sur les médias sociaux, utiliser une application de santé, etc.), et sont traitées et structurées pour alimenter les systèmes algorithmiques.
- 2 La place centrale que les données occupent désormais dans nos sociétés se reflète non seulement dans les discours publics et médiatiques, mais aussi dans les travaux de recherche, alors qu'un nombre croissant d'études s'intéresse de manière spécifique aux implications de leur production, de leur accumulation et de leurs usages, dans tous les domaines de la vie sociale, de la santé à l'économie, en passant par la politique, l'éducation ou la culture. Parmi ces travaux, un ensemble d'études critiques émerge depuis une décennie, qu'on regroupe parfois sous l'étiquette de *critical data studies* (Kitchin et Lauriault, 2014 ; Dalton, Taylor et Thatcher, 2016 ; Hepp, Jarke et Kramp, 2022a ; Kitchin, 2025), que l'on peut traduire en français par « études critiques sur les données » et qui se situent à la croisée de la communication, de la sociologie et des *science and technology studies* (STS). Ce champ de recherche interdisciplinaire est en pleine expansion dans l'espace francophone, où de plus en plus de chercheuses et de chercheurs s'intéressent à la *datafication* et, plus particulièrement, aux relations entre les données numériques et la société. Le grand nombre de propositions reçues dans le cadre de l'appel du présent numéro et la variété des thématiques et terrains empiriques couverts par les articles publiés témoignent de la vitalité de ce champ en contexte francophone.

- 3 Ces travaux se sont développés d'abord comme une réponse aux visions utopiques formulées dans les récits des grandes entreprises technologiques qui dominent l'économie numérique, où le processus de *datafication* de la société est présenté comme étant inéluctable, nécessaire et même souhaitable. Apparue au début des années 2010, la *datafication* — ou « mise en données », en français — renvoie à la fois à la transformation des activités et des sujets sociaux en *data* et à une logique d'exploitation de ces données à des fins économiques et politiques (Mayer-Schönberger et Cukier, 2013 ; Southerton, 2022). Rapidement, les études se sont cristallisées autour de critiques importantes à l'encontre des acteurs de la *Big Tech* et du *Big Data* en particulier (boyd et Crawford, 2012 ; Mejias et Couldry, 2024). Les nouvelles possibilités offertes par les données à l'échelle de la société (par exemple, le mouvement autour des données ouvertes dans les administrations ou dans les sciences) ont été discutées en parallèle des risques en matière de surveillance, de vie privée et de reproduction des inégalités sociales, avec une préoccupation croissante pour les questions éthiques et pour les individus et collectifs représentés dans les données (ou à qui elles sont destinées) (Hepp, Jarke et Kramp, 2022b). Plus récemment, les travaux se réactualisent autour de critiques sur « l'algorithmisation » de la société et, plus généralement, sur le déploiement sous tous azimuts des systèmes d'IA, notamment sur les discours d'accompagnement réaffirmant le caractère inexorable des processus en cours et imposant des visions essentialistes et déterministes des données et de leur production.
- 4 Mais que signifie étudier les données de manière critique ? Il s'agit d'abord de considérer les données comme étant fondamentalement *politiques*, donc comme des instruments de pouvoir (Dalton et Thatcher, 2014 ; Hepp, Jarke et Kramp, 2022a), qui façonnent et qui sont, en même temps, façonnées par des structures sociales, culturelles et économiques (Bigo, Isin et Ruppert, 2019). Il s'agit, ensuite, d'orienter le projecteur non seulement vers la manière dont les données sont produites, utilisées, exploitées, transformées ou détournées en pratique (Millerand *et al.*, 2025), mais aussi vers la façon dont elles circulent dans des espaces discursifs investis d'imaginaires et de valeurs qui tendent à les naturaliser et à les dépolitiser, en les réduisant à leur simple dimension technique (Crawford, Grey et Miltner, 2014). Il s'agit, enfin, de les aborder non pas comme des « données », au sens d'éléments *déjà là*, comme des faits ou des « contenants » de signification non discutables, mais plutôt comme des construits pour en comprendre toutes les ramifications sociopolitiques (Gitelman, 2013 ; Denis, 2018).
- 5 Les textes réunis dans ce numéro interrogent la place centrale des données dans la société contemporaine à partir de trois grands questionnements.
- 6 Un premier ensemble d'articles s'intéresse aux **régimes de production et d'exploitation des données pour saisir de manière critique le processus de *datafication* en cours, dans son macrocontexte**. Ces articles analysent comment les *data* peuvent être utilisées comme des instruments de pouvoir à travers la marchandisation de données relatives aux individus ou leur incorporation dans des stratégies de surveillance et dans des modes de gouvernance étatique et de régulation des conduites.
- 7 L'article d'**Élodie René** et **Samuel Tronçon**, « La gouvernance de la civilisation écologique en Chine : vers une *datafication* de l'intégrité écologique des agents », s'intéresse à la *datafication* en Chine, où celle-ci est au cœur d'une réforme morale de la société. À partir d'une analyse des discours officiels, les auteurs analysent le « système de crédit écologique », qui constitue la branche environnementale du système de crédit

social chinois. Ce crédit écologique mesure et note le niveau d'intégrité écologique des entreprises et des individus afin d'encourager les comportements écoresponsables tout en sanctionnant ceux jugés nuisibles à l'environnement. Au cœur de cette rhétorique verte se jouent à la fois le contrôle des données environnementales au service d'un projet politique autoritaire et la proposition d'un contre-modèle idéologique à l'Occident où fusionnent le « vert » et « l'intelligent ».

- 8 L'article de **Marie-Josée Dupuis, Julien Hocine, Simon Fitzbay et Normand Landry**, « Pratiques numériques et mise en données de la défense des droits humains », aborde quant à lui la *datafication* à l'échelle de la société civile et des organisations engagées dans la défense de droits humains. À partir d'une revue de la littérature sur le sujet sur plus d'une dizaine d'années, les auteurs dégagent un vaste répertoire de pratiques, où les données servent, par exemple, à constituer des preuves de violation des droits, à sensibiliser des publics et à susciter l'action en faveur des victimes, à étayer des contre-récits face à des discours officiels inexacts, etc. Il reste que les groupes de défense des droits humains s'approprient les données et les dispositifs numériques dans un contexte marqué par des rapports de pouvoir asymétriques à l'avantage de leurs adversaires, où les configurations techniques, les cadres réglementaires et la capacité d'action des États — notamment en matière de surveillance et de répression numériques — s'imposent à eux.
- 9 En abordant la *datafication* à l'échelle non pas des États ou de la société civile, mais à celle des municipalités, l'article d'**Étienne Grenier et Nicolas Chartier-Edwards**, « Vie de laboratoire urbain et quartiers intelligents. Analyse des trajectoires de Sidewalk Toronto et du Quartier des spectacles de Montréal » compare deux projets d'infrastructures de type « ville intelligente ». Les auteurs s'intéressent aux données numériques en tant qu'« artefacts politiques » (selon la formule de Winner, 1980), c'est-à-dire comme des sites possibles de controverses susceptibles d'orienter le devenir des projets. Leur analyse des trajectoires du défunt projet de quartier intelligent Sidewalk Toronto et du déploiement massif de dispositifs de captation vidéo dans le Quartier des spectacles à Montréal illustre comment le fait de cultiver l'ambiguïté autour de l'usage des données peut contribuer à diluer les controverses et à neutraliser les débats publics, par exemple en matière de protection de vie privée, et à favoriser ainsi leur acceptabilité sociale.
- 10 Les régimes de production et d'exploitation des données sont aussi analysés à l'échelle des grandes entreprises du numérique. L'article de **Camille Alloing, Ambre Marionneau et Julien Pierre**, « Pour l'amour de la donnée. Vers une *datafication* de nos affects ? », analyse la mise en données des expressions émotionnelles des usagers et usagères par quatre grandes entreprises propriétaires de plateformes de médias sociaux : Google, Meta, ByteDance et X (anciennement Twitter). À partir d'un corpus de publications scientifiques, de brevets et de discours promotionnels, les auteurs montrent comment la transformation des écrits, des interactions et même des corps des usagers et usagères en données « affectives » — qualifiées de sentiments, d'émotions, d'humeurs ou encore de ressentis par les plateformes — constitue un puissant levier de croissance économique. D'une part, ces données affectives reposent sur des représentations universalisantes des émotions qui permettent aux plateformes d'invisibiliser leurs modes de gouvernance ; d'autre part, elles constituent une accapARATION commerciale de « ce qui émeut » en ligne.

- 11 L'article de **Rémi Rouge**, **Emmanuelle Caccamo** et **Katharina Niemeyer**, « Promesses technomnésiques et *datafication* de la mémoire personnelle », s'intéresse lui aussi à la *datafication* à des fins commerciales. L'auteur et les autrices analysent les promesses formulées dans les discours d'entreprises qui développent des technologies de mise en données de la mémoire personnelle, comme Google Photos, Apple Photos, Memoro ou Personal AI. Leur étude illustre empiriquement ce que la *datafication* peut faire à la mémoire. En l'occurrence, ces dispositifs opèrent une sélection normative du mémorable selon des critères qui ne répondent pas tant aux besoins des utilisateurs et utilisatrices, mais plutôt à ceux (techniques, commerciaux, etc.) des entreprises qui les produisent. La promesse d'automatisation des processus mnésiques, telle qu'elle est présentée par ces entreprises, est censée répondre à deux problèmes que ces mêmes entreprises ont contribué à construire, à savoir « le trop-plein de données numériques et la faillibilité de la mémoire humaine ».
- 12 Un deuxième ensemble d'articles s'intéresse à la ***datafication* dans différents secteurs d'activités professionnelles**, pour en examiner les conséquences sur l'évolution des métiers et la redistribution des expertises ou encore sur les pratiques des utilisateurs et utilisatrices de dispositifs basés sur des données collectées sur eux et elles, et ce, dans les domaines de la santé, de la culture, des médias et de l'ouverture des données publiques administratives.
- 13 L'article de **Manon Plegat**, « La *datafication* des soins en maisons de santé : un pouvoir professionnel inégalement distribué », et celui de **Sylvie Grosjean** et **Fabienne Martin-Juchat**, « *Datafication* d'un corps faillible et pratiques d'autosoins. Explorer les *data valences* pour penser la sur-veillance », s'intéressent tous deux aux données de santé. Le premier questionne le décalage entre les objectifs de la mise en données des soins et les usages effectifs qui sont faits des données informatisées. À partir du cas français des maisons de santé, l'autrice montre que les systèmes d'information censés favoriser la coordination pluriprofessionnelle autour des patients ne sont utilisés en pratique que par les médecins généralistes — en partie parce qu'ils reposent sur une conception curative de la santé. Le travail des professionnels non-médecins se trouve quant à lui invisibilisé au fil du processus de mise en machine, de mise en données et de mise en chiffres de la coordination pluriprofessionnelle. Le deuxième article s'intéresse à la manière dont les données de santé générées par des technologies algorithmiques de télésuivi peuvent soutenir des pratiques d'autosoins chez les patients atteints de maladies chroniques. Se basant sur deux études de cas, un dispositif utilisé par des personnes atteintes de la maladie de Parkinson et un implant de *monitoring* pour l'insuffisance cardiaque, les autrices décrivent différents modes de « sur-veillance », où les patients, les dispositifs technologiques et les professionnels de santé coconstruisent une vigilance partagée, porteuse d'encapacitation. À l'opposé d'un modèle de surveillance reposant sur le contrôle et la mesure, la *datafication* ouvre ici un espace dialogique où les données agissent comme les médiateurs d'un apprentissage chez les patients pour apprendre à vivre avec un corps faillible et imprévisible.
- 14 Dans un autre domaine, celui des médias, mais toujours sur la question de la transformation des métiers et des expertises, l'article de **Cassandra Burnier**, « La *datafication* des médias. De la main des journalistes à la mécanique de la donnée », se penche sur la redistribution de l'autorité éditoriale entre les journalistes et les systèmes algorithmiques. Se basant sur l'ethnographie d'un projet d'automatisation d'une infolettre dans un groupe de presse en Belgique, l'autrice montre que l'autorité

éditoriale, traditionnellement incarnée par les journalistes, se transforme en une compétence distribuée entre des acteurs humains (rédacteur en chef, travailleurs de la donnée) et autres-qu'humains (algorithme, métriques). Ces résultats nuancent l'idée d'une dépossession du pouvoir journalistique au profit de la machine, en montrant plutôt un enjeu de délimitation des frontières entre une « logique éditoriale artisanale (centrée sur le jugement, la qualité et la responsabilité de la curation) » et une « logique mécanique de la donnée (centrée sur la performance, la mesure et l'efficacité marketing) ».

- 15 L'article de **Cécile Rodrigues** et **Thomas Chevallier**, « Donnée administrative, ouvre-toi ! Le travail scientifique sur des données de subventions associatives », s'intéresse quant à lui à l'ouverture des données administratives et à leur appropriation. Les données administratives, notamment celles concernant les financements publics, sont longtemps restées « fermées » ou difficiles d'accès. En outre, lorsqu'elles sont « ouvertes », c'est-à-dire mises à disposition en ligne, elles le sont selon une logique de l'offre, et non de la demande, ce qui explique en bonne partie pourquoi elles ne sont pas utilisées par les organisations et la société civile. À la suite d'une démarche entreprise dans le cadre d'un projet de recherche pour accéder à des données sur les subventions fournies par une municipalité au milieu associatif, les auteurs montrent comment les chercheurs et chercheuses peuvent jouer un rôle d'intermédiaire au service d'une politique d'*open data* véritablement orientée vers les usagers et usagères des données. À l'inverse d'une logique de l'offre, il s'agit pour les producteurs de données d'ouvrir un dialogue avec les réutilisateurs potentiels afin de « coconstruire les données ouvertes », en fonction des préoccupations et contraintes de chacun.
- 16 Enfin, l'article de **Juliette Parmentier**, « Cultures de la quantification et quantification de la culture. Les effets de la mise en nombre des pratiques littéraires sur les utilisateurs de plateformes numériques », questionne les effets de la quantification lorsqu'elle s'applique à une activité aussi socialement valorisée que les pratiques littéraires. Se basant sur des entretiens avec des lecteurs qui produisent ou consultent du contenu littéraire sur des applications comme Babelio, Livraddict ou Instagram, l'autrice avance que les « cultures de la quantification » propres à ces dispositifs favorisent « une inflation des pratiques littéraires et un rapport managérial au soi lecteur ». Elle pointe les contradictions entre les discours (une conception noble de la lecture axée sur le plaisir) et les pratiques (poussant les lecteurs à faire croître leurs métriques). Ainsi, à la différence d'autres pratiques, dont la quantification de soi au service de l'activité sportive, la quantification des pratiques littéraires n'implique pas seulement de maximiser ses métriques : « le bon lecteur lit *plus*, et *assidûment*, mais ne lit pas *trop* — ce qui serait perçu comme un rapport déplacé, compétitif ou consumériste à la littérature ».
- 17 Le troisième ensemble d'articles traite de questions plus générales sur la *datafication* et les données, qui amènent à réfléchir aux **enjeux de littératie des données, aux modalités pratiques de la production de données, et à la place du travail humain en particulier, ainsi qu'aux études critiques sur les données comme champ d'études**.
- 18 L'article d'**Ugo Verdi**, « Lutter contre les violences des données issues de la *datafication*. Pour une littératie des données *queer* », pose les bases théoriques d'une littératie des données, pensée comme un outil de lutte contre les effets néfastes de la mise en données sur les vies des personnes *queer*. L'auteur part du constat que ces personnes

subissent une triple violence dans le processus de *datafication* en cours : une invisibilisation du fait de données manquantes sur leurs réalités, une essentialisation de leurs caractéristiques du fait de biais persistants (contribuant par exemple à la production de profils *queer* caricaturaux par les systèmes algorithmiques) et une stigmatisation liée à l'usage de données les concernant à des fins de surveillance (par exemple au sein des applications de rencontres). L'article propose de définir une littérature *queer* centrée sur une approche quintuple — critique, « mécanologique », *queer*, intersectionnelle et éthique — qui permette la participation active aux débats sur la mise en données des vies *queer* et, ultimement, le développement de méthodes de conception, d'analyse et de représentation des données *queer*.

- 19 L'article de **François Lambotte** et **Kévin Carillon**, « L'artisanat des données à l'ère de la *datafication* », introduit et développe le concept d'artisanat des données pour qualifier un ensemble de pratiques qui participent à la production des données. À partir de deux enquêtes ethnographiques, l'une menée dans des organismes responsables de la production de registres du cancer et l'autre dans la communauté Stack Overflow sur Stack Exchange, un forum en ligne fondé par des développeurs de logiciels, les auteurs montrent que, dans certains contextes, la production de données s'apparente à un travail d'artisanat impliquant des savoir-faire, une certaine éthique du soin et des dynamiques communautaires. Ces pratiques révèlent la dimension fondamentalement humaine du travail des données, un travail minutieux et collectif, certes laborieux et souvent invisibilisé, mais qui mobilise des formes d'expertise, d'engagement collectif et de sens chez les travailleuses et travailleurs concernés, à l'opposé des représentations dominantes dénonçant des pratiques largement automatisées, décontextualisées et désincarnées.
- 20 Pour conclure ce numéro, l'article de **Florence Millerand**, **Lucie Delias** et **Mélanie Millette**, « Les études critiques sur les données. Généalogie et cartographie d'un champ », propose une cartographie du champ d'études en retraçant son émergence au sein de la littérature, principalement anglophone. Les autrices montrent comment il a émergé au croisement de différentes filiations théoriques, en réponse au positivisme naïf des études sur le *big data* qui dominaient les recherches au début des années 2010. Rassemblées autour d'un objet de recherche propre, la « donnée », et d'une épistémologie sociotechnique et critique, les études critiques sur les données proposent une approche transversale des enjeux sociaux, politiques et économiques liés aux technologies numériques, dont l'exploitation des données est à la base du fonctionnement. L'article synthétise trois apports majeurs à l'étude des technologies en sciences humaines et sociales. D'abord, les données ne sont ni neutres ni brutes, mais situées, contextuelles et traversées de rapports de pouvoir (où certains producteurs de données occupent une position dominante dans leur production et leur exploitation à grande échelle). Ensuite, la *datafication* de la société vient reproduire et intensifier des logiques préexistantes aux technologies numériques, de quantification et d'automatisation notamment. Enfin, les enjeux autour de la gouvernance et de la souveraineté des données dans le contexte actuel marqué par le développement rapide de l'IA, d'un côté, et des velléités politiques autour de la préservation ou de la destruction de jeux de données desservant certaines stratégies politiques dominantes, de l'autre, rappellent la dimension éminemment politique des données et leur intérêt stratégique.

- 21 Ensemble, les textes rassemblés dans ce numéro offrent une vue sur plusieurs des problématiques qui animent actuellement le champ des études critiques sur les données, présentant une diversité de terrains empiriques et ouvrant de nouvelles perspectives d'études pour saisir les mutations en cours et à venir.

BIBLIOGRAPHIE

- BIGO Didier, Engin ISIN et Evelyn RUPPERT (2019), *Data politics. Worlds, subjects, rights*, Londres, Routledge.
- BOYD danah et Kate CRAWFORD (2012), « Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon », *Information, Communication & Society*, 15(5), p. 662-679.
- COULDRY Nick et Ulises A. MEJIAS (2024), *Data grab. The new colonialism of Big Tech and how to fight back*, Londres, Penguin.
- CRAWFORD Kate, Mary L. GRAY et Kate MILTNER (2014), « Critiquing big data: Politics, ethics, epistemology », *International Journal of Communication*, 8(0), p. 1663-1672.
- DALTON Craig et Jim THATCHER (2014), « What does a critical data studies look like, and why do we care? Seven points for a critical approach to "Big Data" », *Society & Space*, <https://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care>, page consultée le 30 novembre 2025.
- DALTON Craig, Linnet TAYLOR et Jim THATCHER (2016), « Critical data studies: A dialog on data and space », *Big Data & Society*, 3(1), <https://journals.sagepub.com/doi/10.1177/2053951716648346>, page consultée le 5 décembre 2025.
- DENIS Jérôme (2018), *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines.
- GITELMAN Lisa (dir.) (2013), « Raw data » is an oxymoron, Cambridge, MIT Press.
- HEPP Andreas, Juliane JARKE et Leif KRAMP (2022a), « Introduction », dans Andreas HEPP, Juliane JARKE et Leif KRAMP (dir.) (2022), *New perspectives in critical data studies: The ambivalences of data power*, Londres, Palgrave MacMillan, p. 1-23.
- HEPP Andreas, Juliane JARKE et Leif KRAMP (dir.) (2022b), *New perspectives in critical data studies: The ambivalences of data power*, Londres, Palgrave MacMillan.
- KITCHIN Rob (2025), *Critical data studies: An A to Z guide to concepts and methods*, Cambridge, Polity Press.
- KITCHIN Rob et Tracy LAURIAULT (2014), « Towards critical data studies: Charting and unpacking data assemblages and their work », dans Jim THATCHER, Andrew SHEARS et Josef ECKERT (dir.), *Thinking big data in geography: New regimes, new research*, Lincoln, University of Nebraska Press.
- MAYER-SCHÖNBERGER Viktor et Kenneth CUKIER (2013), *Big data: A revolution that will transform how we live, work, and think*, Boston, Houghton Mifflin Harcourt.

MILLERAND Florence, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (2025), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l'Université de Montréal.

SOUTHERTON Clare (2022), « Datafication », dans Laurie A. SCHINTLER et Connie L. MCNEELY (dir.), *Encyclopedia of big data*, New York, Springer International Publishing, p. 358-361.

AUTEURS

FLORENCE MILLERAND

Florence Millerand est professeure titulaire au Département de communication sociale et publique de l'Université du Québec à Montréal. Courriel : millerand.florence@uqam.ca

LUCIE DELIAS

Lucie Delias est maître de conférences en sciences de l'information et de la communication au département Infocom de l'Université Paul-Valéry Montpellier 3. Courriel : lucie.delias@univ-montp3.fr

MÉLANIE MILLETTE

Mélanie Millette est professeure titulaire au Département de communication sociale et publique de l'Université du Québec à Montréal. Courriel : millette.melanie@uqam.ca

Dossier

Articles

La gouvernance de la « civilisation écologique » en Chine

Vers une datafication de l'intégrité écologique des entreprises et des individus

Élodie René et Samuel Tronçon

- 1 Au tournant des années 2000-2015, la Chine est confrontée à une crise environnementale sans précédent, menaçant non seulement l'intégrité de son territoire et la santé de ses habitants, mais également la légitimité du Parti Communiste Chinois (PCC) (Albert et Xu, 2016). La pollution de l'air, de l'eau et des sols, exacerbée par des décennies d'industrialisation rapide et de croissance économique effrénée, atteint des niveaux critiques. Les scandales liés à la falsification de données environnementales et les violations répétées des lois par les entreprises d'État, souvent avec la complicité des gouvernements locaux, ont alimenté une méfiance croissante envers l'appareil d'État (Monjon et Poncet, 2018). À l'international, les images des « airpocalypses » — d'intenses épisodes de pollution — dans des villes comme Pékin ou Chongqing ont terni l'image de la Chine et renforcé les critiques sur son modèle de développement (Mouton, 2022).
- 2 Au moment de son arrivée au pouvoir en 2012, Xi Jinping hérite d'un paysage institutionnel fragmenté où les autorités locales contournent régulièrement les lois nationales en matière de protection sociale et environnementale. La falsification et la distorsion des données est monnaie courante, limitant la capacité du gouvernement central à évaluer les performances des cadres locaux (Huang, 2020). Ces pratiques exacerbent la crise de légitimité que traverse alors le PCC (Hammer, 2017) et érodent la confiance publique dans l'appareil d'État. C'est dans ce contexte que Xi Jinping érige la protection environnementale parmi les priorités stratégiques nationales et fait du concept de Civilisation Écologique un des slogans phares du Parti (Sullivan, 2023).
- 3 Le Président chinois va entamer une refonte ambitieuse du rôle de l'État dans la gouvernance environnementale du pays. L'objectif affiché est de mobiliser l'innovation numérique — intelligence artificielle (IA), méga données, et plateformes connectées — pour optimiser la gestion de la crise environnementale à l'aide d'outils de

quantification des performances socio-environnementales des agents économiques. Ce projet de modernisation écologique du pays s'inscrit dans une logique sécuritaire où les mathématiques souveraines deviennent un levier politique pour gouverner par les nombres (Weigelin-Schwiedrzik, 2018). Si cette approche n'est pas spécifique à la Chine, reflétant en réalité un processus global propre au développement des États modernes (Dérosières et Kott, 2005 ; Foucault, 2025/1972, 2007/1979 ; Lascoumes et Le Galès, 2018 ; Scott, 1998 ; Supiot, 2015), elle prend ici une dimension singulière en raison du contrôle autoritaire exercé par le PCC sur les systèmes d'information du pays.

- 4 Cet article explore comment les ambitions du PCC en matière de contrôle des données environnementales se manifestent à la fois sur le plan discursif, à travers les récits du Parti sur la Civilisation Écologique Numérique (CEN), et sur le plan opérationnel, via le Système de Crédit Écologique (SCE). Encore peu connu en Occident, le SCE correspond à la branche environnementale du Système de Crédit Social chinois. Il vise à mesurer et à noter le niveau « d'intégrité écologique » des entreprises et des individus afin d'encourager les comportements écoresponsables tout en sanctionnant ceux jugés nuisibles à l'environnement (Fang, 2022¹). Deux sous-systèmes du SCE ont été ici étudiés : le système de crédit environnemental des entreprises et le système de crédit écologique des citoyens, actuellement en phase d'expérimentation à Lishui. Cette analyse permet d'interroger les fondements idéologiques du système de crédit écologique ainsi que son efficacité opérationnelle pour atteindre les objectifs climatiques fixés par Pékin : un pic des émissions d'ici 2030 et la neutralité carbone à l'horizon 2060. Les concepts clés mobilisés par les chercheurs chinois travaillant à la conceptualisation de la civilisation écologique et numérique et du SCE ont ainsi été étudiés au prisme des *Critical Data Studies*. Dans cette perspective, une dizaine de textes séminaux sur la CEN et le SCE ont été analysés. Ces textes émanent d'« intellectuels organiques » au sens de Gramsci (1971/1929-1935), c'est-à-dire de cadres et d'experts du Parti qui contribuent à élaborer et stabiliser les doctrines écologiques du PCC tout en les articulant à la pensée de Xi Jinping. La sélection et l'échantillonnage des textes a été faite à partir de trois principaux critères : l'autorité institutionnelle et l'expertise des auteurs et contributeurs clés, la temporalité et la contemporanéité des textes, et enfin leur référencement sur le web chinois. Les textes retenus reflètent les productions académiques, idéologiques et doctrinales de chercheurs chinois et d'experts institutionnels occupant des positions clés dans les organismes de recherche ou de pilotage de la CEN et du SCE en Chine. Ils sont soit directement rédigés par ces auteurs, soit issus d'entretiens ou de comptes rendus de réunions publiques où ceux-ci prenaient la parole. Ils traitent de la construction institutionnelle et idéologique du SCE ainsi que de la vision du développement promue par la CEN, illustrant différentes dimensions (normative, opérationnelle, idéologique) du modèle chinois d'écologie politique. L'ensemble des textes étudiés ont été publiés entre 2020 et 2024 et font l'objet de reprise de contenu importante sur le web chinois. Les écrits étudiés proviennent de plateformes reconnues (sites gouvernementaux et médias officiels) ce qui permet d'appréhender le matériau comme représentatif des discours de légitimation et de pilotage institutionnel de la CEN.

La civilisation écologique numérique

La civilisation écologique : une rhétorique verte au service d'un projet politique autoritaire

- 5 Le concept de Civilisation Écologique a émergé, au cours des deux dernières décennies, comme un pilier central de la rhétorique environnementale du Parti Communiste Chinois pour faire face à la crise de légitimité auquel il était alors confronté (Marquis *et al.*, 2011). Ce concept repose sur l'idée d'une coexistence harmonieuse entre l'humanité et la nature, promue comme une vision alternative au développement durable tel qu'il est conceptualisé dans les pays occidentaux (Marinelli, 2018). Désormais intégré au socle idéologique du Parti-État, il est présenté comme un cadre théorique et pratique offrant une perspective holistique dans laquelle l'homme et la nature cohabitent symbiotiquement (Gaffric et Heurtebise, 2013).
- 6 Cependant, derrière cette apparente quête d'harmonie civilisationnelle, la civilisation écologique constitue avant tout un outil de politique stratégique. Elle sert à renforcer la légitimité du PCC tout en proposant un contre-modèle idéologique à l'Occident, souvent décrit dans les discours officiels comme prédateur et impérialiste. En mobilisant un héritage confucéen idéalisé, ce concept vise à promouvoir un modèle chinois distinct, en opposition aux valeurs capitalistes occidentales (Hansen, 2018 ; Heurtebise, 2023). Par ailleurs, la civilisation écologique s'accompagne d'une rhétorique de modernisation écologique qui associe innovation technologique, protection environnementale et croissance économique pour consolider l'autorité du Parti (Goron, 2018).
- 7 Avec l'arrivée au pouvoir de Xi Jinping en 2012, le contrôle de l'État sur les discours relatifs à la civilisation écologique s'est intensifié de manière significative (Buckley, 2021 ; Geall et Ely, 2018 ; Goron, 2018 ; Hansen, Li et Svarverud, 2018 ; Wang, 2022). Ce concept est devenu l'un des slogans clés de la rhétorique du Parti, accompagnée d'une production prolifique de discours et de politiques sur le sujet. Xi Jinping a promu, à tous les niveaux de la société, une vision triomphante du renouveau écologique de la Chine dirigée par le Parti (Goron, 2018). Une vision qui prône un modèle descendant d'écologie politique, visant à répondre par la force et l'autorité aux attentes des citoyens chinois face aux scandales environnementaux et à l'impunité des pollueurs. En affichant une volonté ferme de lutter contre la pollution et d'améliorer la transparence environnementale, le PCC de Xi Jinping cherche à regagner la confiance populaire tout en renforçant son emprise sur les entreprises, administrations locales, et les citoyens (Marquis et Bird, 2018). Ce positionnement stratégique permet également au PCC de présenter la civilisation écologique comme un modèle global capable d'inspirer d'autres nations dans leur transition écologique (Li et Shapiro, 2020).

La civilisation écologique numérique : une évolution stratégique

- 8 Avec la prolifération et la complexification des discours sur la civilisation écologique sous Xi Jinping, le PCC a décliné ce concept en plusieurs versions. Parmi elles, la Civilisation Écologique Numérique (CEN) illustre l'ambition du PCC d'articuler transition écologique et transformation numérique comme leviers stratégiques de modernisation du pays.

- 9 Selon les intellectuels organiques du PCC, la CEN repose sur deux dynamiques convergentes : la « numérisation de la civilisation » et l'« écologisation de la civilisation », qui se rencontrent pour transformer les relations entre nature, société et technologie (Shi et Chi, 2024). Présentée comme une nouvelle étape historique après les civilisations primitives, agricoles et industrielles, cette vision ambitionne de transposer les dimensions matérielles, spirituelles et institutionnelles de la civilisation écologique dans des formes numériques afin d'en optimiser la gestion. Les technologies numériques sont ainsi mobilisées pour améliorer la gouvernance environnementale, rationaliser l'utilisation des ressources naturelles et renforcer la productivité économique tout en favorisant une coexistence harmonieuse entre l'homme et son environnement (*ibid.*).
- 10 Parmi les objectifs stratégiques de la CEN, on retrouve la volonté de mettre en place un système intégré de collecte, d'analyse et de partage des données environnementales à travers des infrastructures numériques avancées. Il s'agit de promouvoir la création d'un centre national de données écologiques, conçu comme un *data hub* permettant de coordonner les efforts interrégionaux en matière de gouvernance environnementale (Huang et Luo, 2024). Un tel système s'appuierait sur des technologies telles que la surveillance par satellite, les drones intelligents ou encore les capteurs IoT (*Internet of Things*), permettant une collecte automatisée des données relatives à l'atmosphère, au sol ou aux écosystèmes marins (Baidu Encyclopedia, 2023).
- 11 Shi et Chi (2024), politologues à l'école du Marxisme de l'Université de Fujian dans le Zhejiang, insistent sur le rôle clé de la *datafication* dans l'optimisation des processus décisionnels liés à l'écologie (Shi et Chi, 2024). Par exemple, il est question d'utiliser des jumeaux numériques (*digital twins*), comme celui développé pour le bassin fluvial Yongjiang dans le Zhejiang, afin de simuler des scénarios environnementaux complexes tels que l'évolution des inondations. Ces outils pourraient permettre non seulement une gestion proactive, mais aussi une anticipation des risques écologiques grâce à des modélisations prédictives basées sur des ensembles massifs de données. Au-delà des écosystèmes naturels, ce processus de *datafication* englobe également les comportements humains via des mécanismes de contrôle écologiques individuels et collectifs (Huang et Luo, 2024). Ces dispositifs sont censés inciter les citoyens et les entreprises à adopter des pratiques écoresponsables grâce à un suivi algorithmique de leurs activités quotidiennes. Il s'agit de développer des dispositifs de surveillance numérique des conduites pour écologiser les comportements individuels et collectifs.
- 12 Cette approche traduit un projet de gestion centralisée de l'écologie fondé sur les données, visant à englober aussi bien les entreprises que les écosystèmes naturels et les comportements individuels au sein d'un cadre unifié. En mobilisant un réseau intégré d'infrastructures numériques, la CEN a l'ambition de transformer chaque élément naturel ou social en entité mesurable et traçable (*ibid.*).
- 13 Elle s'inscrit ainsi dans la continuité des grands récits politiques portés par Xi Jinping depuis 2012, valorisant l'usage stratégique des outils numériques pour renforcer le contrôle du PCC sur la société et l'économie (Wang, 2023). Les discours sur la CEN font écho à ceux sur la grande réjuvenation de la nation chinoise, qui présentent l'innovation technologique comme un moteur de modernisation institutionnelle et de croissance économique dans les secteurs de pointe (Geall et Ely, 2018 ; Gourdon, Monjon et Poncet, 2016). Dans cette vision, le numérique devient à la fois un levier d'efficacité environnementale et un instrument de contrôle politique au service du

projet civilisationnel d'État. Xi Jinping a ainsi transformé la doctrine du PCC sur la civilisation écologique, initialement conçue comme un cadre d'analyse des crises du modèle de développement chinois, en un récit optimiste de régénération écologique piloté par le numérique (Goron, 2018). Cette rhétorique légitime le renforcement des capacités de surveillance verticale de l'État au nom de la modernisation la gouvernance environnementale du pays et de la lutte contre l'impunité des pollueurs (Zhou, 2020).

Le numérique au service d'une écologie centralisée

- 14 En 2014, le Comité permanent de l'Assemblée nationale populaire a adopté une version révisée de la loi sur la protection de l'environnement, sous l'initiative de Xi Jinping. Celle-ci introduit des normes plus strictes pour contrôler les pollutions tout en renforçant les obligations de transparence environnementale des entreprises d'État. Elle prévoit également la mise en place de dispositifs d'évaluation des performances écologiques des cadres locaux du PCC pour les contraindre à faire appliquer les lois environnementales nationales. Ces mesures visent à renforcer les capacités de surveillance des autorités centrales sur les échelons administratifs et politiques inférieurs afin de responsabiliser les gouvernements locaux et de mieux contrôler leurs actions (Ker et Logan, 2014). Deux ans plus tard, en 2016, le gouvernement central a réformé le système d'application des lois environnementales du pays afin de renforcer le pouvoir exécutif du ministère de l'Écologie et de l'Environnement (MEE). Cette réforme vise officiellement à redonner au MEE le pouvoir de superviser directement les bureaux locaux de protection environnementale, qui étaient jusqu'alors sous la direction des gouvernements provinciaux ou municipaux (Zhou, 2020). L'objectif affiché est de limiter la capacité de ces derniers à contourner les lois nationales (*ibid.*). Naturellement, de telles réformes impliquent une stratégie de contrôle accru de l'information environnementale par les autorités centrales.
- 15 Parallèlement, Xi Jinping a renforcé le contrôle de l'État sur les plateformes numériques environnementales ouvertes au public — telles que WeChat ou les microblogs des bureaux locaux du MEE. Officiellement ces dispositifs sont censés encourager la participation du public à la gouvernance informationnelle de l'environnement (Zhang, Mol et He, 2016). Ils permettent aux citoyens d'accéder à des données horaires sur la qualité de l'air ou de l'eau via leurs téléphones intelligents et offrent aussi des canaux pour signaler des problèmes environnementaux, notamment à travers des applications mobiles développées par certaines organisations non gouvernementales environnementales chinoises (Kostka, Zhang et Shin, 2019). Bien que ces initiatives digitales soient officiellement présentées comme faisant partie intégrante d'une stratégie nationale de modernisation écologique attentive aux préoccupations des citoyens, elles incarnent en réalité un modèle autoritaire de gouvernance environnementale où l'État exerce un contrôle quasi absolu sur la production et la diffusion des informations environnementales (Bolsover et Goron, 2020 ; Monjon et René, 2021).
- 16 La transparence promue, en apparence, par ces systèmes doit ainsi être interprétée avec prudence. Derrière l'ouverture affichée se cache un contrôle strict du discours environnemental en ligne. Les pratiques telles que la falsification des données ou l'inondation des plateformes avec du contenu homogène reprenant le discours officiel ont été documentées (Goron et Bolsover, 2020 ; Zeng *et al.*, 2017). Toute initiative

citoyenne indépendante qui dépasse le cadre institutionnel contrôlé par l'État est sévèrement réprimée (Fu et Distelhorst, 2018). Cette répression s'est largement renforcée sous Xi Jinping et s'accompagne même désormais d'une criminalisation systématique des activistes environnementaux qui contreviennent aux objectifs du PCC, ce qui n'était pas le cas sous Hu Jintao (*ibid.*).

- 17 Dans cette vision autoritaire et techno-centrée de l'écologie, la production de l'information environnementales doit rester sous le contrôle des autorités centrales et être conforme à l'idéologie du Parti. Si ce dernier prône l'avènement d'une nouvelle ère de développement vert « axée sur l'innovation » (Geall et Ely, 2018), ce projet ne laisse pas de place aux voix dissidentes ou à d'autres conceptions de l'écologie. La civilisation écologique et numérique promue par le Parti-État allie donc autoritarisme vert et contrôle numérique, via une vision de l'innovation technologique qui ne se limite plus à la seule sphère technique, mais englobe également des expérimentations politiques et institutionnelles. L'objectif est d'intégrer les technologies numériques dans tous les aspects de la société, de la gouvernance environnementale au contrôle des modes de vie des citoyens, en passant par la gestion optimisée des entreprises et la maîtrise des risques écologiques. En s'appuyant sur des systèmes de contrôle de plus en plus sophistiqués, le Parti souhaite aligner les comportements des citoyens, des entreprises et des cadres du Parti avec ses objectifs civilisationnels (Monjon et René, 2021). Depuis quelques années, cette stratégie s'est concrétisée à travers la mise en place de divers dispositifs de contrôle numérique de la fiabilité écologique des agents qui sont expérimentés à l'échelle nationale à travers différents projets pilotes locaux.

Opérationnaliser la confiance sociale

Le système de crédit écologique

- 18 Le système de crédit écologique (SCE) incarne pleinement cette articulation entre modernité technologique, nationalisme culturel et contrôle centralisé, dont la CEN prétend faire la synthèse. Ce vaste système aux multiples ramifications illustre de façon concrète la manière dont le PCC mobilise les technologies numériques pour surveiller, évaluer et orienter les comportements des acteurs politiques, économiques et sociaux. L'objectif phare du SCE consiste à mobiliser les technologies digitales les plus avancées pour faire appliquer les lois environnementales du pays sur l'ensemble du territoire, tout en promouvant une « moralité écologique » parmi les citoyens, les entreprises et les cadres du Parti (Fang, 2022).
- 19 Concrètement, le SCE se déploie à travers la mise en place d'une série de dispositifs numériques expérimentaux conçus pour mesurer la « fiabilité écologique » des agents et rationaliser l'usage des ressources naturelles. Ces systèmes visent à sanctionner ceux qui ne respecteraient pas les lois et réglementations environnementales nationales tout en récompensant ceux qui démontreraient un haut niveau de conformité.
- 20 De tels dispositifs visent également à intégrer les externalités socio-environnementales des activités productives dans les coûts de production des entreprises afin de réorienter l'économie chinoise vers des modèles de valorisation fondés sur l'écologie plutôt que sur la finance (Fang, 2022 ; Wu, 2023). Par exemple, les entreprises ayant une faible note écologique peuvent être exclues des marchés publics ou se voir refuser l'accès à certains financements verts. À l'inverse, celles ayant une bonne note

bénéficient d'avantages, tels que des taux préférentiels ou un accès renforcé aux marchés internationaux. Ces systèmes visent à rendre obligatoire la divulgation d'informations environnementales par les entreprises (Tan, 2024) afin de diriger les flux financiers vers les technologies bas-carbone (Wu, 2023).

- 21 Le SCE cherche, par ailleurs, à façonner les subjectivités en promouvant une culture écologique nationale et en orientant les comportements individuels vers l'écoresponsabilité, selon les critères du Parti (*ibid.*). Dans cette perspective, les citoyens chinois sont présentés comme des acteurs clés de la transition écologique devant être activement mobilisés pour construire une nouvelle culture écologique nationale. À travers des plateformes numériques dédiées, ces derniers sont encouragés à signaler les comportements polluants ou à documenter leur empreinte carbone personnelle via des applications interactives. Officiellement, il s'agit d'instaurer une gouvernance écologique fondée sur la participation collective et la co-responsabilité entre l'État, les entreprises et la société civile (Fang, 2022).
- 22 Le SCE revêt donc une dimension civilisationnelle, en tant qu'outil opérationnel visant à écologiser — via le contrôle informationnel — les modes de gouvernance et de production du pays, mais aussi les modes de vie des citoyens. Il est d'ailleurs décrit comme un outil qui permettra à la Chine, une fois les expérimentations achevées, de mettre en place une société nouvelle fusionnant le « vert et l'intelligent ». Pour bien comprendre son fonctionnement et ses fondements idéologiques, il est indispensable de le situer dans le contexte plus large du système de crédit social chinois.

Le SCE comme sous-système du système de Crédit Social (SCS)

- 23 Encore peu connu en occident, le SCE correspond en réalité à la branche environnementale du Système de Crédit Social (SCS) chinois. Lancé en 2014 par Xi Jinping, le SCS s'inscrit dans une démarche de modernisation de la gouvernance et de moralisation sociale à travers le déploiement de nouveaux outils de surveillance et de notation des agents (Creemers, 2018). Initialement conçu comme un mécanisme de notation financière visant à réduire les risques de défaut de paiement, le SCS s'est progressivement élargi pour inclure la régulation des comportements sociaux, économiques et environnementaux. Ce vaste système repose sur un ensemble complexe de mécanismes numériques, plus ou moins sophistiqués, permettant de collecter, analyser et partager des données issues d'une variété d'acteurs publics et privés. Ces données sont utilisées pour attribuer des notes aux individus, entreprises et institutions en fonction de leur conformité aux lois nationales et aux normes morales promues par le Parti (Sel, 2019).
- 24 L'objectif affiché du SCS est d'instaurer une « société de haute confiance » en renforçant l'intégrité dans les affaires gouvernementales, commerciales et sociales, ainsi que la crédibilité judiciaire (Chen, Lin et Liu, 2018). Ce système ambitionne également de combler les lacunes dans l'application des lois et de résoudre les problèmes liés à la distorsion des données (Von Blomberg, 2020). Les comportements jugés « dignes de confiance » sont récompensés par des incitations économiques ou administratives tandis que ceux considérés comme « indignes » entraînent des sanctions allant jusqu'à l'exclusion sociale ou économique (Cheung et Chen, 2024). Les institutions financières utilisent désormais les données générées par le SCS pour

ajuster les taux d'intérêt ou refuser des services aux individus ou aux personnes morales jugés « non fiables » (*ibid.*).

- 25 Ce système comprend ainsi deux facettes : un appareil normatif encourageant les « bons » agents et la moralité sociale et économique ; et un appareil disciplinaire punissant les comportements « déviants » et « indignes de confiance » (Zhang, 2020). L'ambition du SCS est de mettre en place une constellation de mécanismes automatisés de sanctions combinées, où les personnes physiques et morales figurant sur les listes noires (*trust-breaking*) sont soumises à des restrictions multiples sans procédure indépendante. Ces sanctions peuvent inclure l'interdiction de voyager, des contraintes administratives, économiques ou sociales, et même des impacts sur les membres de la famille des individus concernés (Cheung et Chen, 2024).
- 26 Ce vaste réseau repose sur quatre mécanismes principaux : la collecte massive de données comportementales via un réseau institutionnel complexe ; le partage d'informations entre différentes plateformes centralisées telles que *Credit China* ; un système de labellisation permettant d'attribuer des statuts tels que « fiable » ou « non fiable » ; et enfin un mécanisme coordonné de sanctions appliquées par plusieurs agences gouvernementales pour maximiser leur effet dissuasif (Liang *et al.*, 2018). À travers ces fonctions interdépendantes, le SCS incarne une transition vers un modèle de gouvernance où la *datafication* devient l'instrument principal du contrôle social.
- 27 L'ambition officielle de Xi Jinping, initiateur du projet, est de mettre en place un modèle de gouvernance basé sur la quantification des comportements humains collectifs et individuels. Il s'agit de construire une société guidée par les données (*data-driven society*), où les comportements humains sont constamment surveillés par l'État (Liang *et al.*, 2018). Les données collectées sont utilisées pour créer une identité numérique (*data self*) qui tend à dominer l'identité biologique (*bio-self*), renforçant ainsi la dépendance des individus et des entreprises envers l'État (Cheung et Chen, 2024 ; Drinhausen et Brussee, 2021). La notion de *data self* est centrale dans ce modèle, l'identité numérique des individus et des personnes morales devenant déterminante pour les opportunités offertes ou les contraintes qui s'imposent dans la vie réelle.
- 28 Notons toutefois qu'il est important de relativiser le discours officiel chinois, l'analyse de ces systèmes révélant une réalité assez éloignée d'une parfaite numérisation de l'ensemble de la société, instrumentée par de puissants outils numériques. Contrairement à certaines perceptions occidentales, le système chinois de crédit social reste partiellement numérisé et repose encore largement sur des enquêtes humaines et des décisions locales. Cela laisse place à des disparités dans l'application des sanctions entre les niveaux local et central (Drinhausen et Brussee, 2021). De plus, les entreprises restent les principales cibles du système, tandis que la notation individuelle est encore marginale et souvent volontaire (*ibid.*).

SCE : un système aux fondements idéologiques hybrides

- 29 Le système de crédit écologique élargit les fonctions du système de crédit social pour répondre aux exigences de la civilisation écologique numérique. Alors que le SCS se concentre principalement sur la quantification et l'évaluation des interactions humaines dans les sphères sociale et économique, le SCE intègre également les interactions entre l'humain et son environnement naturel.

- 30 Sa conception s'appuie sur une hybridation entre des mécanismes généralement identifiés comme néolibéraux, tels que l'évaluation numérique et l'incitation par le marché, des outils hérités de la période soviétique et maoïste, comme le contrôle centralisé du Parti-État, et des discours civilisateurs glorifiant l'héritage culturel chinois. Plutôt que d'introduire une rupture philosophique ou un basculement idéologique, les discours relatifs à la civilisation écologique et numérique et au Système de crédit écologique s'appuient sur une synthèse pragmatique de différentes idéologies politiques. On retrouve d'ailleurs, dans les productions intellectuelles chinoises sur le SCE, plusieurs références à Friedrich Hayek (1960). Ce dernier est mobilisé pour sa vision d'une gestion rationnelle par le marché des enjeux environnementaux, où l'intégration du crédit écologique dans les mécanismes financiers repose sur une logique incitative visant à encourager des comportements écovertueux. Garret Hardin (1968) est également fréquemment cité pour sa conception de la tragédie des biens communs. Il est mobilisé afin de justifier la nécessité d'une régulation stricte des écosystèmes naturels pour prévenir leur surexploitation, ce qui conduit à préconiser l'articulation entre des mécanismes de marché (la privatisation étant présentée comme plus efficace que la gestion par les communs pour protéger les espaces naturels) et des régulations étatiques autoritaires, appuyées sur des outils numériques permettant de contrôler les usages (Fang, 2022). Ces références cohabitent avec celles sur le socialisme aux caractéristiques chinoises et sur le confucianisme.
- 31 Ainsi, la civilisation écologique et numérique du PCC et les systèmes de gouvernance associés sont emblématiques de ce que Ngai-Ling Sum décrit comme un « mode autoritaire de régime ordolibéral-socialiste » (Sum, 2019). En Chine, la marchandisation et la forte intervention de l'État vont de pair : le marché joue un rôle clé dans l'allocation des ressources, mais reste étroitement contrôlé par l'appareil étatique. Contrairement au « laissez-faire » prôné par l'école de Chicago, l'ordolibéralisme chinois préconise une intervention étatique forte pour structurer la société selon les logiques du marché. Ce modèle favorise une organisation sociale centrée sur la croissance du PIB où individualisme, concurrence et innovation sont encouragés par un environnement normatif façonné par l'État.
- 32 Ce syncrétisme idéologique trouve sa cohérence dans le recours généralisée à la quantification socioécologique comme outil clé pour structurer les interactions entre l'État, le marché, les citoyens et la nature dans un cadre autoritaire. À cet égard, les *Critical Data Studies* offrent une grille de lecture intéressante pour bien comprendre la façon dont le SCE est mobilisé par le PCC pour instrumenter son pouvoir ordo-libéral au nom de la civilisation écologique et numérique. Cette littérature met en évidence la façon dont les données, loin d'être neutres, sont produites, interprétées et utilisées pour structurer les relations de pouvoir au sein des États modernes. Comme le souligne Scott (1999), la quête de lisibilité étatique — c'est-à-dire la capacité de l'État à rendre ses sujets et ses ressources visibles et contrôlables — est une caractéristique historique des projets modernes de gouvernance (Foucault, 2007 ; 1977). Dans cette logique, la mise en données des comportements humains devient un levier central pour gouverner par les nombres (Supiot, 2015). Comme le montrent Espeland et Vannebo (2007), la statistique — née comme une « science de l'État » (Hacking, 1990) — ne se limite pas à décrire la réalité sociale : elle contribue activement à la produire en classant, hiérarchisant et orientant les pratiques. Le SCE s'inscrit pleinement dans cette dynamique : il transforme les performances sociales et environnementales des agents

en données mesurables, permettant à l'Parti-État de surveiller, évaluer et orienter leurs comportements pour les rendre conformes à son projet de civilisation écologique.

- 33 En prétendant mobiliser des infrastructures numériques sophistiquées pour instrumenter les comportements des acteurs économiques et sociaux, le déploiement du SCE reflète non seulement l'ambition du PCC d'asseoir son contrôle autoritaire sur l'ensemble de la société, mais également les tensions inhérentes à ce modèle hybridant un contrôle étatique centralisé avec des mécanismes de marché valorisant la compétition et l'esprit d'entreprise. Comme le montrent Bourdieu (2018) et Loveman (2005), la manière dont l'État collecte, traite et met en forme les données pour représenter la réalité sociale porte une forte charge symbolique. Ce travail de représentation ne se contente pas d'organiser la société, mais contribue également à définir qui a accès aux ressources publiques et aux opportunités sociales (Fourcade, 2021 ; Johns, 2021). Dans cette perspective, la quantification de la fiabilité écologique peut donc être interprétée comme une nouvelle dimension du contrôle social, offrant un levier puissant pour récompenser ou punir certains acteurs sociaux en fonction d'une norme socialement acceptable (la protection de l'environnement), mais dont la définition est totalement verrouillée par l'appareil idéologique du Parti.
- 34 Ainsi, le recours aux *Critical Data Studies* permet d'éclairer deux dimensions essentielles du SCE : d'une part, son rôle dans la production d'une société lisible et gouvernable par le biais des données ; d'autre part, son inscription dans un projet plus large visant à redéfinir les rapports entre économie, société et écologie dans une perspective civilisationnelle. Le SCE illustre bien comment le PCC utilise la *datafication* non seulement comme un outil technique, mais aussi comme un instrument idéologique pour légitimer son autorité tout en façonnant un modèle unique de gouvernance environnementale. Il s'agit à présent, pour en évaluer l'efficacité réelle, d'examiner de plus près les mécanismes de notation, de sanction et de récompense implémentés dans ces systèmes, tout en interrogeant les logiques sous-jacentes à cette quantification : quels comportements sont valorisés ou sanctionnés ? Comment ces processus participent-ils au renforcement des structures de pouvoir existantes ?

La *datafication* au cœur d'une réforme morale de la société

- 35 Pour cette analyse, nous avons exclusivement utilisé des données secondaires accessibles en ligne sur le web chinois. Ces données ont été collectées à partir de moteurs de recherche chinois et des sites web officiels de structures gouvernementales. Les sources principales incluent des rapports officiels, des lignes directrices en matière de politiques publiques, ainsi que des textes techniques et réglementaires explicitant le fonctionnement SCE et ses sous-systèmes, publiés par diverses agences nationales et provinciales impliquées dans sa mise en œuvre.
- 36 Bien que l'approche méthodologique repose principalement sur une analyse qualitative approfondie, nous avons également collecté des données quantitatives provenant de plateformes officielles, notamment dans la province du Zhejiang. Ces données consistent principalement en des listes de classement d'entreprises, consultables en ligne, et qu'il s'agit de récupérer pour effectuer des comptages et analyser la distribution des notations dans la population. En complément, des informations ont été

extraites du site officiel d'Alibaba, acteur privé pionnier dans le développement de systèmes de crédit écologique. En raison des contraintes géographiques et des restrictions d'accès à certains interlocuteurs chinois, aucune donnée primaire n'a été recueillie via des entretiens ou d'autres méthodes directes.

Le crédit écologique des entreprises

- 37 Le système de crédit environnementale des entreprises (SCEE) prend la forme d'un système national hiérarchisé et centralisé — le *Cadre national sur l'évaluation du crédit environnemental des entreprises* (MEE, 2014) —, qui reste néanmoins très hétérogène tant par le format des données recueillies que par les critères analysés. Il comprend un mouvement ascendant, traversant les différentes échelles administratives depuis les organisations locales jusqu'à l'État central, et qui assure la captation, principalement en auto-déclaration, la vérification et la complétion des données. Un mouvement inverse, assure le pilotage du système, tant dans la définition des objectifs et des critères, que dans leur contrôle opérationnel, avec un pilotage multi-échelle du système. À cette hiérarchisation territoriale est associée une multiplicité de systèmes territoriaux, respectant des lignes directrices nationales, déclinées régionalement, voire à l'échelle d'une agglomération.
- 38 Cette multiplicité n'est pas contradictoire avec l'ambition d'un système totalement centralisé, et semble au contraire jouer un rôle dans le dessein global. En effet, la littérature technique (*ibid.*) sur le SCEE témoigne d'une volonté d'innover par l'expérimentation (China Business Network, 2024) afin de favoriser l'évolution rapide des systèmes de notation vers des formes stables et ayant fait leurs preuves sur le terrain. Le caractère multi-échelle et territorialisé de ce système de sont des facteurs puissants de variabilité contrôlée. Ils peuvent donc permettre une sélectivité progressive des critères les plus efficaces et les plus robustes au moment de la numérisation et de l'agrégation. Par ailleurs, l'analyse de la méthodologie des dispositifs de notation montre que le système global est justement conçu de manière à regrouper, à faire coexister, voire à faire converger des méthodes, des proxys et des mesures fortement hétérogènes².
- 39 Au-delà de cette nature expérimentale, la variabilité des systèmes est aussi le marqueur de fortes disparités. C'est à l'échelon régional que sont identifiés et définis les principaux indicateurs prenant en compte les caractéristiques physiques, démographiques et économiques. Cette deuxième variabilité semble acceptée par l'administration centrale, le système étant partiellement tributaire de ces disparités du fait de la définition des proxys en fonction de caractéristiques locales : la présence d'une industrie imposante ou d'une agriculture extensive, l'ancienneté et le niveau de transformation de l'outil de production, changent, de fait, la typologie même des mesures qui peuvent être établies, et donc une partie des résultats.
- 40 Enfin, une troisième variabilité concerne les techniques, modes de collecte et niveaux d'intégration numérique des territoires. Il y a une corrélation assez évidente entre le niveau de sophistication des systèmes territoriaux et leur puissance économique. Mais on peut supposer aussi que, derrière l'appropriation plus ou moins élevée du projet d'ensemble, se cachent des niveaux de criticalité ou des enjeux politiques disparates, y incluant notamment la diligence plus ou moins affirmée des cadres locaux eu égard à un projet désigné comme particulièrement stratégique au plus haut niveau de l'État. Le

projet d'une administration numérique qui permettrait de gommer les disparités en mettant tous les citoyens et toutes les organisations à égalité semble donc pour l'instant buter sur les mêmes inégalités que dans le monde réel.

- 41 La province du Zhejiang est désignée par les autorités centrales comme le principal laboratoire d'innovation de Chine (Wang, 2019). De plus, son système de crédit écologique est actuellement le plus sophistiqué et le plus opérationnel, ce qui en fait un excellent candidat pour notre étude de cas. La notation prend la forme d'une échelle de points à partir de laquelle on déduit un rang. Un référentiel d'évaluation liste les critères thématiques auxquels correspondent un nombre maximal de points pouvant être attribués en fonction du niveau de réalisation du critère. Chaque critère comprend un certain nombre d'indicateurs qui vont permettre d'enlever ou ajouter des points. Les indicateurs peuvent être associés à des proxys ou des mesures empiriques dont l'objectivité est précisée dans un cahier des charges technique. Pour chaque critère, on va trouver une gamme de niveaux d'atteinte permettant de se représenter la note de l'entreprise sur ce critère, indépendamment des modalités d'évaluation empiriques. Ainsi, le *Zhejiang Enterprise Environmental Credit Evaluation Management Measures* pourra par exemple annoncer que l'entreprise respecte les attentes réglementaires, est exemplaire, ou n'est pas conforme aux attentes réglementaires sur le critère analysé (MEE, 2014).
- 42 En cumulant les différents critères, qui ont chacun un poids différent, on arrive à une note sur 1 000, qui permet de classer les entreprises. Le classement comporte deux niveaux d'excellence, deux niveaux d'insuffisance et un niveau médian. L'idée étant que les niveaux d'excellence ouvrent l'accès à des dispositifs d'encouragement, de promotion et de facilitation afin d'inciter les entreprises à respecter les critères. En particulier, les entreprises classées dans le plus haut niveau auront une priorité dans l'accès au crédit et à l'investissement, aux licences et brevets, ainsi qu'à tout un ensemble de dispositifs pour assurer leur promotion et leur réputation. Tandis que les entreprises classées dans les niveaux d'insuffisance seront limitées sur ces ressources, voire totalement exclues, notamment concernant les marchés publics (MEE, 2014).
- 43 Le simple fait de limiter l'accès au crédit est déjà une sanction dans la mesure où une entreprise a besoin de financer le cycle de trésorerie, sans compter des investissements éventuels pour développer ses marchés. On peut donc aisément imaginer que, dès le niveau médian, l'activité économique de la plupart des entreprises soit déjà soumise à des contraintes importantes.
- 44 Concernant le niveau le plus bas, il est difficile de savoir si les entreprises sont explicitement poussées au dépôt de bilan, le croisement des données avec le registre des entreprises n'étant pas facilité. Mais le référentiel mentionne explicitement le fait que les non-conformités ayant déclassé une entreprise dans les niveaux inférieurs doivent être corrigées dans l'année si l'entreprise veut espérer un reclassement positif. Mais dans le cas contraire, il va de soi que l'absence d'outils de trésorerie au-delà d'un an, même avec un modèle économique robuste, signe quasiment l'arrêt de toute activité ou la dégradation rapide de l'outil productif.
- 45 On ne connaît pas bien les conditions réelles d'application de ces mesures, mais le référentiel mentionne explicitement une gradation importante en fonction des infractions qui pourrait laisser penser que seules les infractions répétées ou extrêmes, c'est-à-dire mettant en danger la sécurité des personnes ou dégradant l'environnement sur des échelles importantes, sont susceptibles d'entraîner l'ajout de l'entreprise sur

une liste noire. Pour les autres infractions, le référentiel montre au contraire une relative « tolérance » pour récupérer un niveau correct à condition de démontrer la transparence des déclarations et la diligence à corriger les erreurs ou payer les contraventions. Cette réversibilité reste néanmoins à l'interprétation des organismes de contrôle et de pilotage ainsi que, en principe, de la société civile (médias, organisations non gouvernementales, population).

Un modèle en rupture avec la vision occidentale de l'impact environnemental

- 46 La particularité de ce système de crédit environnemental des entreprises réside aussi dans la manière dont sont produites et administrées les données. S'il s'inspire ouvertement des systèmes de pointage financier existant en Europe et aux États-Unis, on peut aussi le comparer, au moins du fait des dimensions analysées, aux initiatives européennes d'évaluation de l'impact social et environnemental des entreprises.
- 47 Le pointage financier est en général interne aux institutions bancaires, les clients sont évalués, mais sans avoir un accès libre à leur note. Les référentiels de calcul ne sont pas transparents et les algorithmes ne sont pas ouverts (Bücker *et al.*, 2022). Par ailleurs, l'accès aux données n'est pas une pratique développée, malgré l'existence du Règlement général de protection des données (RGPD) qui institue un droit d'accès (Union européenne, 2016). Si l'accès aux données des transactions constitue déjà une source d'information dérogatoire au respect de la vie privée, l'étendue des données qui sont ou peuvent être croisées grâce au méga données par les organismes de pointage financier reste inconnue (Berg *et al.*, 2020).
- 48 À l'opposé, les systèmes d'évaluation de la responsabilité sociétale des entreprises sont renseignés volontairement par celles-ci. Des standards d'évaluation ont été développés, fonctionnant souvent comme des labels, qui permettent aux entreprises d'assurer une forme de « transparence » et de comparabilité de leur démarche d'amélioration sociale et environnementale (Stamm, 2021). Néanmoins, ces démarches sont souvent sectorielles ou affinitaires, et de nombreuses entreprises choisissent de concevoir elles-mêmes leur modèle d'évaluation. Enfin, la CSRD (*Corporate Sustainability Reporting Directive*), modèle européen unifié de publication d'informations sur la durabilité, n'est pour l'instant obligatoire que pour les grandes entreprises.
- 49 En comparaison, le SCEE chinois combine des mécanismes volontaires et obligatoires. Comme l'indiquent les documents-cadres nationaux, les évaluations sont ouvertes et transparentes. Certaines typologies d'entreprises (notamment les industries lourdes) sont tenues de participer à l'évaluation, tandis que les autres sont encouragées à s'y soumettre de manière volontaire (MEE, 2014).
- 50 Les méthodes de pointage sont publiques, et les entreprises, ainsi que toutes leurs parties prenantes, ont accès aux notes finales des sites audités (MEE, 2014). De plus, les données sont partagées par toute la chaîne hiérarchique du SCEE. Elles peuvent être agrégées pour former des tableaux de bord à différentes échelles territoriales. On est donc dans un modèle de partage de la donnée stratégique qui emprunte à dessein plusieurs des caractéristiques techniques des systèmes occidentaux, tout en évitant leur opacité ou leur relativité apparente pour en faire un argument d'objectivité et donc d'« adoptabilité ».

- 51 Dans un tel système, l'enjeu de la sincérité des données est central. Le référentiel prévoit donc des contrôles aléatoires sur échantillonnage ainsi qu'un système d'audits qui sont réalisés par des organismes tiers accrédités et soumis au SCEE. Les tailles d'échantillon sont définies pour chaque rang, avec des dimensions d'autant plus importantes que la notation est basse. Plus les entreprises ont une note élevée, moins les contrôles sont fréquents, et inversement. Il faut noter que le fait de moins contrôler les entreprises ayant une notation élevée est un avantage de plus offert aux meilleurs niveaux de notation, avantage qui peut paraître incohérent avec l'objectif de lutte contre l'insincérité des déclarations et qui peut engendrer une reproduction des inégalités entre acteurs.
- 52 L'analyse de la structure du calcul du SCEE du Zhejiang révèle que divers types de critères sont regroupés sous les catégories « social » et « écologique » (MEE, 2014). Les critères dits « environnementaux » regroupent la protection écologique dans l'utilisation des ressources, dans la sélection des sites industriels et dans le développement et la construction. Ils incluent également les différentes catégories de pollution : pollution de l'air, de l'eau, déchets solides, gestion des eaux usées, déchets dangereux, pollution sonore... Ces différentes dimensions environnementales ne sont pas quantifiées en tant que telles, leur évaluation n'est pas construite comme une variable dans un modèle productif écosystémique. Elles sont analysées uniquement sous l'angle du respect ou non des objectifs réglementaires, nationaux et régionaux, en fonction de la filière ou de la production. En somme, les critères écologiques mesurent le respect des normes fixées par l'État, plutôt que la compatibilité du modèle économique avec des objectifs de protection de l'environnement. Un deuxième groupe de critères dits « sociaux » porte sur le rapport à la société civile, exigeant une bonne réputation médiatique et l'absence de pétitions ou de plaintes citoyennes. Ces critères établissent un contrôle exercé par les médias et la population qui semble contradictoire avec le rôle limité accordé aux médias et à la parole citoyenne. Trois autres critères concernent la gestion de l'information sur la pollution et les processus de production, qui doivent être gérés de manière transparente et fiable. Ces critères ne font pas référence à la responsabilité sociale des entreprises, mais au contrôle social des externalités négatives environnementales : ils évaluent la transparence des données transmises au système vu comme un outil de contrôle pour et par la société, garanti par l'État. Le critère social repose donc sur la confiance de la société dans l'appareil de production et sur la préservation de l'ordre public.
- 53 Ainsi, l'État central, avec ce système de contrôle généralisé des comportements, dispose d'un réseau d'acteurs économiques prêts à adapter leur production et leur impact en fonction des injonctions politiques, et à fournir des données en temps réel permettant aux pouvoirs publics de piloter la « réponse » écologique. Une telle approche évalue les politiques publiques uniquement du point de vue de la réponse sociale aux injonctions publiques, plutôt qu'en fonction de l'effet attendu sur l'environnement. Nous sommes donc bien face à une vision autoritaire et verticale des politiques écologiques. Toutefois, le SCEE pris comme système d'évaluation peut sembler relativement « neutre » du point de vue de la structure sociopolitique nécessaire à son fonctionnement. Même s'il est présenté comme un outil visant à moderniser l'économie socialiste de marché en Chine, sa conception et l'accent mis sur la transparence écologique peuvent avoir pour effet de le rendre apparemment compatible avec les économies libérales de marché.

- 54 Le SCEE semble donc avoir une double vocation, comme outil visant à piloter les politiques écologiques en Chine, et comme « format » exportable à l'international, adaptable à différents contextes politiques tout en diffusant une vision chinoise de l'écologie politique.

Une efficacité relative révélatrice d'une double fonction opérative-transformative

- 55 On peut s'interroger sur l'efficacité du modèle de gouvernance informationnelle autoritaire, souvent perçu comme un exemple de capitalisme de surveillance dirigé par l'État (Huang et Tsai, 2022), où l'information joue un rôle central pour maintenir une vision descendante de l'écologie. La surproduction de données sur les performances environnementales peut créer une illusion de contrôle (Huang, 2020). Ces dernières années, les données environnementales publiées par le Parti-État ont mis en avant les progrès remarquables de la Chine en matière de protection de l'environnement et de réduction des émissions de CO₂ (FMPRC, 2022). Cependant, ces publications officielles doivent toujours être appréhendées avec prudence. La production des données environnementales reste strictement contrôlée par l'État qui sélectionne, en définitive, quelles informations peuvent être divulguées et quels indicateurs doivent être utilisés pour surveiller la pollution (Sun et Huang, 2021 ; Yu et Yan, 2020). De plus, malgré les efforts de Xi Jinping pour lutter contre la corruption et renforcer le contrôle de l'information, la falsification des données et les pratiques de corruption demeurent importantes, notamment aux échelons locaux (Huang, 2020). Il est donc impossible de conclure à un impact environnemental positif du système de crédit écologique, notamment lorsqu'on regarde de plus près le système de crédit environnemental des entreprises. En outre, bien que les discours officiels affirment une volonté d'un développement vert plus qualitatif, la priorité centrale du Parti reste la croissance économique et la relance de la demande intérieure comme l'affirme le dernier plan quinquennal chinois (2026-2030). Dans ce contexte, la protection de l'environnement ne saurait entraver les objectifs de relance économique portés par les autorités chinoises (Huang, 2021).
- 56 Il en ressort que les différents outils de contrôle déployés dans le cadre du système de crédit écologique servent avant tout le discours du PCC. Ainsi, les données publiques « démontrent » l'excellence des entreprises en matière de conformité environnementale. En analysant dans le détail les chiffres du système régional de crédit environnemental des entreprises au Zhejiang, on découvre que plus de 90 % des sites industriels contrôlés entre janvier et février 2021³ étaient conformes à la grande majorité des indicateurs utilisés dans le système d'évaluation. Ces chiffres varient légèrement chaque année, mais restent valides en 2025, malgré l'apparente rigueur du système de notation. Cette notation globalement très positive semble incohérente avec les nombreux rapports publiés par des organisations internationales (Wateratrisk, Greenpeace, etc.), dont il est intéressant de noter qu'ils datent tous d'avant 2017 et qu'il est depuis difficile de trouver des données indépendantes actualisées (Greenpeace, 2014 ; Hu, 2017). Un rapide aperçu des autres provinces semble indiquer la même répartition des niveaux de conformité. On peut interpréter ce résultat soit comme l'indice d'une forme de tolérance des institutions locales à l'égard de chiffres manipulés voire falsifiés, soit comme le résultat de normes insuffisantes à limiter la pollution et la

destruction des écosystèmes. On peut également émettre l'hypothèse que l'aspect fortement composite des critères d'évaluation engendre une déperdition de l'efficacité des outils de mesure et donc de leurs effets sur les transformations comportementales.

- 57 Il n'en reste pas moins que le nombre d'entreprises effectivement sanctionnées par ce système reste très limité, et l'autoritarisme vert (Gilley, 2012) paraît relativement complaisant avec les industriels qui contribuent le plus à la puissance économique du pays.
- 58 Au-delà des questions d'efficacité et d'efficience du système de crédit écologique et des sous-systèmes associés comme le système de crédit environnemental des entreprises, se pose donc la question du rôle qu'ils jouent dans le projet global de Civilisation Écologique et Numérique. On constate notamment que la dépolitisation des enjeux écologiques à travers les mécanismes de *pointage* prend tout son sens dans les projets de finance verte que la Chine entend promouvoir. La Banque Populaire de Chine, le Ministère de l'Écologie et de l'Environnement, l'administration d'État de la Supervision Financière et la Commission de Régulation des Valeurs Mobilières de Chine élaborent progressivement une doctrine visant à renforcer la finance verte comme levier de la transition écologique. Cette doctrine anticipe notamment l'intégration des performances des acteurs économiques en matière de réduction du carbone, de diminution de la pollution et de croissance verte dans les critères d'octroi de prêts, de calcul des quotas et de fixation des taux d'intérêt (China Business Network, 2024).
- 59 Plusieurs défis opérationnels sont identifiés par ces acteurs, qui attestent tout à la fois des ambitions de co-intégration du système financier à la politique environnementale, et d'une immaturité technique de l'infrastructure de crédit écologique à ce stade.
- 60 C'est d'abord un enjeu de conceptualisation des constantes socio-environnementales et de leur mesure. Ainsi, l'absence de normes unifiées pour évaluer les comportements écologiques complique la quantification des contributions environnementales. La pénurie de compétences spécialisées oppose également un frein sérieux, notamment pour la financiarisation des variables écologiques et l'ingénierie des mesures d'impact. D'autant plus que l'évaluation porte aussi sur des « comportements écologiques » et non uniquement sur des mesures physiques, ce qui accentue la complexité de la démarche (China Business Network, 2024).
- 61 C'est ensuite un enjeu de modélisation systémique des organisations et des flux sensibles au contexte. Toute la gamme des technologies est évidemment convoquée, des méga données à l'Internet des objets, en passant l'IA. Pourtant, au-delà de la technologie, c'est la source même de la quantification qui présente un véritable défi : complexité des externalités à mesurer, contextes physiques de prélèvement et de transformation des ressources naturelles, grandes étapes des processus industriels, jonctions entre acteurs clés des filières... Autrement dit, tous les « lieux » du système productif où peuvent arriver des erreurs, des fraudes, des fuites et des impacts non maîtrisés.
- 62 Pour les acteurs du système financier chinois, la financiarisation des impacts environnementaux des activités humaines reste donc un objectif qui nécessite d'importants investissements matériels et immatériels pour arriver à pleine maturité. Ce constat ne fait que renforcer l'hypothèse selon laquelle le système de crédit écologique n'est pas une fin en soi, mais constitue un dispositif de transition vers la civilisation écologique et numérique. Actuellement en phase d'initialisation, le SCE

semble surtout destiné à instaurer un régime de data-surveillance, tout en travaillant l'acculturation des acteurs, les boucles de renforcement et l'acceptabilité sociale du système. Dans une phase ultérieure, une fois que les freins de conceptualisation et de modélisation seraient levés, il ne resterait plus qu'à déployer massivement l'exploitation opérationnelle des données et des formats de *datafication* adaptés à de très nombreux contextes de mesure sociale et environnementale.

Vers un crédit écologique des citoyens

- 63 Si pour l'instant, le système de crédit écologique concerne principalement les entreprises, l'État teste aussi des systèmes de notation pour évaluer les comportements des citoyens (Fang, 2022). Une fois de plus, la province du Zhejiang est à la pointe de l'expérimentation à Lishui, ville-pilote de la construction d'une civilisation écologique en Chine (Wang, 2019), et zone de démonstration nationale pour la mise en place d'un « système de crédit social de haute qualité » intégrant l'écologie dans l'évaluation de l'intégrité sociétale des individus (Suzhou.gov.cn, 2021). Dans cette optique, les autorités locales ont publié une liste des comportements écologiques positifs et négatifs servant de base à la mise en place d'un système local de crédit écologique des citoyens (Lishui City, 2020). Cette liste comprend une série de critères à partir desquels peut être évalué le caractère plus ou moins écologique du mode de vie des individus (déplacements, consommation, déchets, économies d'énergie, etc.), mais aussi de la culture locale (bien-être public, sacrifice civilisé, funérailles vertes, etc.) (*ibid.*).
- 64 Dans sa mise en application, ce système s'appuie sur plusieurs sources de données. Par exemple, pour les résidents de Lishui, la note personnelle résulte d'un calcul combinant à parts égales la note de crédit social provincial (géré à l'échelle du Zhejiang) et la note de crédit écologique attribuée au niveau municipal. Le dispositif repose sur une plateforme d'information, qui centralise et met à jour en temps réel les données collectées par les différentes administrations locales. Le suivi et la gestion du processus relèvent du bureau municipal du crédit, qui évalue les comportements écologiques des individus à partir des informations enregistrées sur cette plateforme (Lishui credit, lishui.gov.cn).
- 65 Chaque résident permanent de la ville se voit ainsi attribuer une note correspondant à l'un des cinq grades définis : AAA et AA pour les excellents profils, A pour les comportements jugés satisfaisants, B pour un niveau moyen et C pour les profils présentant des manquements. Les valeurs sont ajustées en fonction de la distribution des résultats dans la population, et un mécanisme garantit la cohérence des calculs entre l'échelle municipale et l'échelle régionale. Par ailleurs, certaines administrations publiques et organisations autorisées par la loi peuvent accéder à ces notes, tandis que la consultation par des tiers est soumise à l'obtention du consentement préalable de la personne concernée.
- 66 En fonction du grade obtenu, les citoyens peuvent obtenir des récompenses : trajet gratuit, prêt sans dépôt dans une bibliothèque publique, accès facilité aux services de soins aux personnes âgées, bon d'achat pour des produits de luxe importés, bons pour des vacances dans des lieux réputés... Ces récompenses s'inscrivent parfaitement dans la ligne idéologique du système de crédit écologique qui vise à mettre en place des « mécanismes d'incitation orientés vers le marché » pour forger des modes de vie écologiques (Fang, 2022). Il est important de noter que ce système de crédit individuel

ne prévoit pour l'instant aucune sanction pour grades inférieurs, il est présenté comme purement incitatif et reposant sur le volontariat des habitants (Suzhou.gov.cn, 2021).

- 67 En parallèle du système développé par l'État, Alibaba — une société chinoise cotée en bourse — a déployé plusieurs applications pour évaluer les comportements écologiques des individus. En 2016, Ant Financial, filiale d'Alibaba, a lancé *Ant Forest*, une application pour renforcer la conscience écologique des utilisateurs, grâce à une approche gamifiée de l'engagement individuel en faveur de la durabilité (Alibaba Foundation, 2018). Les utilisateurs reçoivent des points s'ils font des choix de vie à faible émission de carbone comme se déplacer en transports en commun, ne pas utiliser de vaisselle jetable, acheter des vêtements de seconde main, etc. Avec ces points, ils peuvent faire pousser des « arbres virtuels », qui seront ensuite convertis en arbres réels plantés dans les territoires sujets à la désertification. Plus d'une centaine de marques ont rejoint *Ant Forest* pour y créer leur propre forêt virtuelle, y compris des multinationales occidentales, telles que Starbucks, Timberland ou Estée Lauder (Business Wire, 2021). En 2019, *Ant Forest* reçoit même le prix « champion de la Terre », une des plus hautes distinctions environnementales des Nations Unies, pour ses actions incitatives favorisant les écogestes au sein de la population chinoise (UNEP, 2019).
- 68 Plus récemment, en 2022, Alibaba a introduit un nouveau dispositif de gestion des émissions baptisé « 88 Carbon Account », conçu pour promouvoir l'écoresponsabilité parmi les utilisateurs de ses plateformes de commerce en ligne, principalement au moyen de bons de réduction (Azila, 2022). Le principe est similaire à *Ant Forest*, les utilisateurs gagnent des points lorsqu'ils adoptent des comportements sobres en carbone, qu'ils peuvent convertir en bons d'achat. Les participants sont également invités à publier des preuves visuelles de leurs « gestes d'éco-guerrier », en téléchargeant des vidéos d'eux en train d'éteindre la lumière quand ils quittent une pièce par exemple (*ibid.*). Par le biais de ce projet, Alibaba prétend contribuer à l'atteinte des objectifs nationaux de neutralité carbone à l'horizon 2060, en ciblant une réduction de 1,5 gigatonne d'émissions au sein de son écosystème numérique d'ici 2035 (Yu, 2022). Selon Yang Lingye, directrice de la stratégie et des opérations environnementales, sociales et de gouvernance d'Alibaba, la finalité du programme consiste à sensibiliser les consommateurs aux comportements à faible empreinte carbone et à encourager l'adoption de choix écologiques dans les pratiques quotidiennes (*ibid.*). D'autres acteurs majeurs du secteur numérique chinois, tels que JD.com et Meituan, ont mis en œuvre des initiatives similaires, témoignant de l'engouement des acteurs privés chinois du numérique pour ce type de systèmes.
- 69 Ces différents systèmes privés de notation verte tendent à individualiser la responsabilité écologique sans remettre en cause ni le consumérisme encouragé par Alibaba ni les structures de production liées à sa plateforme de e-commerce. Les fournisseurs stimulent leurs ventes en affichant l'empreinte carbone de leurs produits, tandis que les clients peuvent consommer davantage s'ils démontrent leur écoresponsabilité selon les critères définis par Alibaba. Ces dispositifs permettent à l'entreprise de produire des statistiques de réduction carbone présentées comme objectives et neutres, censées attester de sa contribution aux objectifs nationaux de réduction des émissions. Toutefois, cette contribution reste largement discutable en l'absence de contrôle démocratique sur les méthodes de calcul de l'empreinte carbone des fournisseurs.

Conclusion

- 70 Comme évoqué précédemment, le SCE a pour objectif d'évaluer et de noter « l'intégrité écologique » des entreprises, des citoyens et des agents publics, en s'appuyant sur un ensemble d'indicateurs quantifiés reflétant le degré de conformité de leurs comportements aux objectifs environnementaux définis par l'État (Monjon et René, 2021). Ce dispositif repose sur une logique d'incitation et de dissuasion : il s'agit, d'une part, de récompenser les pratiques jugées vertueuses sur le plan environnemental, et d'autre part, de pénaliser les comportements considérés comme nuisibles, notamment en restreignant l'accès des acteurs sanctionnés à certaines opportunités économiques et sociales (Fang, 2022). À travers ce système, les autorités chinoises cherchent à généraliser le principe d'écoconditionnalité de l'action publique à l'ensemble de la société, inscrivant ainsi cette démarche dans le cadre plus large de la CEN en conformité avec la pensée de Xi Jinping.
- 71 Le pouvoir central déploie ainsi une stratégie écologique à plusieurs dimensions. Il rationalise et encadre les pratiques des administrations et des entreprises dans une optique de modernisation de la gouvernance environnementale (Li et Lang, 2010 ; Wang, 2019). Il met en place un outil de pilotage des comportements individuels, en transformant la « fiabilité écologique » en enjeu économique (et donc en enjeu de distinction sociale). Il projette à l'international une image de puissance verte, à même de relever les défis de son temps, tant sur le plan intérieur qu'au sein de la « communauté de destin » qu'elle appelle de ses vœux à l'international (Hansen, 2018 ; Heurtebise, 2023).
- 72 Tel un « art d'administrer à distance » (Rose, 1990), le SCE s'efforce de gouverner les conduites et d'amener les acteurs à se réguler en internalisant les normes et les valeurs promues par le régime. En ce sens, les fondements idéologiques du SCE doivent être appréhendés en tenant compte des spécificités du contexte politique et culturel chinois. Au-delà de l'impératif de modernisation écologique et de recherche d'efficacité, la construction de la CEN permet d'aborder sous un angle nouveau le rapport entre État, individu, marché et nature. Elle s'inscrit dans une conception holiste du monde articulée autour du nexus « pouvoir, savoir et nature », et se déploie à partir d'une ingénierie incrémentale dont l'objectif est d'amener la société à s'adapter aux impératifs de l'État.
- 73 Cependant, si la dimension morale et incitative de la rhétorique mérite d'être soulignée, c'est la dimension du contrôle qui doit retenir ici l'attention. Cette dernière confère au SCE un caractère propre, qui la distingue des autres types de démarches incitatives. L'ambition d'influencer, voire de contraindre, le citoyen et les entreprises à adopter un comportement vertueux et à rejeter ce qui ne l'est pas, se traduit dans la pratique par un volontarisme politique omniprésent qui ne peut se comprendre qu'en s'intéressant aux sources intellectuelles mobilisées dans ce processus.
- 74 L'analyse des fondements conceptuels du système permet ainsi de mettre en évidence des références théoriques hétérogènes, dont l'articulation permet de saisir les spécificités de la pensée chinoise en matière de gouvernance écologique, tout en sortant de la dialectique Orient/Occident :
- D'abord, une lecture renouvelée de Franz Hayek, qui conduit à confier au marché la double fonction de régulateur de l'action humaine et de révélateur de « bonnes » pratiques. Le

couplage entre notations et système de bonus-malus incite dès lors les agents économiques à intégrer, dans leurs actions, les objectifs de l'État en matière de protection de l'environnement (Fang, 2022).

- Ensuite, une relecture assumée de Garrett Hardin, qui tend à justifier un encadrement renforcé des ressources communes pour éviter leur surexploitation. L'expérimentation du crédit écologique s'appuie sur une théorie renouvelée du contrôle de l'usage des ressources, où les technologies numériques, adossées à des mécanismes de marché, permettent d'évaluer et d'orienter les conduites des agents pour assurer une gestion soutenable des écosystèmes (*ibid.*).
- 75 On peut ainsi voir se dessiner une certaine cohérence idéologique avec les références à la « dictature souveraine » de Carl Schmitt (1921), convoquée dans les textes doctrinaires du PCC qui visent à légitimer un régime où totalitarisme, ordre constitutionnel et dispositif juridique sont pleinement subordonnés au pouvoir politique. Comme le souligne Harvey (2021), malgré la revendication d'une filiation marxiste-socialiste, les références à Mao, Lénine ou Marx sont désormais marginales dans le discours idéologique du Parti sous Xi Jinping, au profit de lectures et d'allusions plus fréquentes à Carl Schmitt.
- 76 En articulant de manière systémique des instruments économiques, des dispositifs technologiques, des outils idéologiques et une ambition de contrôle social, le SCE s'inscrit dans une vision du monde qui dépasse largement la simple sphère environnementale. La question se pose dès lors de savoir si ce dispositif relève avant tout d'un autoritarisme qui intègre l'écologie ou d'une réelle écologie autoritaire. En déléguant à des systèmes la tâche de déterminer quels sont les comportements éocompatibles, assiste-t-on à l'émergence d'une nouvelle écologie politique, ou à l'adaptation du totalitarisme au contexte de stress écosystémique ? La *datafication* des comportements n'opère-t-elle pas une distinction entre « amis » et « ennemis » de la civilisation écologique numérique, et donc de l'État ? Deux questions qui font de la CEN un projet politique dont l'analyse et l'observation seront capitales dans les décennies à venir.

BIBLIOGRAPHIE

ALBERT Eleanor et Beina XU (2016), « China's environmental crisis », *Council on Foreign Relations*, 18, <https://www.cfr.org/backgroundunder/chinas-environmental-crisis>, page consultée le 23 novembre 2025.

BAIDU (2023), « Digital ecological civilization », *Baidu Encyclopedia*, <https://baike.baidu.com/item/%E6%95%B0%E5%AD%97%E7%94%9F%E6%80%81%E6%96%87%E6%98%8E/62954608#reference-2>, page consultée le 23 novembre 2025.

BERG Tobias, Valentin BURG, Ana GOMBOVIĆ et Manju PURI (2020), « On the rise of FinTechs: Credit scoring using digital footprints », *The Review of Financial Studies*, 33(7), p. 2845-2897.

- BOURDIEU Pierre (2018), *On the State: Lectures at the Collège de France, 1989-1992*, Hoboken, John Wiley & Sons.
- BÜCKER Michael, Gero SZEPANNEK, Alicja GOSIEWSKA et Przemyslaw BIECEK (2022), « Transparency, auditability, and explainability of machine learning models in credit scoring », *Journal of the Operational Research Society*, 73(1), p. 70-90.
- BUCKLEY Lila (2021), *Engaging with China's ecological civilisation*, rapport de la Green Economy Coalition, février.
- BUSINESS WIRE (2021), « Over 600 Million People Planted More Than 326 Million Trees via Alipay Ant Forest in Five Years », *Business Wire*, 25 août, <https://www.businesswire.com/news/home/20210825005340/en/Over-600-Million-People-Planted-More-Than-326-Million-Trees-via-Alipay-Ant-Forest-in-Five-Years>, page consultée le 23 novembre 2025.
- CHEN Yu-Jie, Ching-Fu LIN et Han-Wei LIU (2018), « “Rule of Trust”: The Power and perils of China's social credit megaproject », *Columbia Journal of Asian Law*, 32(1), p. 1-36, DOI : <https://doi.org/10.7916/CJAL.V32I1.3369>.
- CHEUNG Anne S. et Yongxi CHEN (2022) « From datafication to data state: Making sense of China's social credit system and its implications », *Law & Social Inquiry*, 47(4), p. 1137-1171.
- CHINA BUSINESS NETWORK (2024), « Ecological credit enters the policy toolbox to determine the “gold content” with “green content” », *Sina*, octobre, <https://finance.sina.com.cn/money/smj/2024-10-19/doc-inctpeuc7629712.shtml>, page consultée le 23 novembre 2025.
- CRÉDIT SUZHOU (2021), « Lishui Zhejiang: Fully launching the “Ecological Credit” brand », *Crédit Suzhou*, Septembre, <https://credit.suzhou.gov.cn/news/show/377033.html>.
- CREEMERS Rogier (2018), « China's social credit system: An Evolving practice of control », *SSRN Electronic Journal*, Janvier, DOI : <https://doi.org/10.2139/ssrn.3175792>.
- DESROSIÈRES Alain et Sandrine KOTT (2005), « Quantifier », *Genèses*, 1, p. 2-3.
- DRINHAUSEN Katja et Vincent BRUSSEE (2021), « China's social credit system in 2021: From fragmentation towards integration », *Merics China Monitor*, 3 mars, p. 1-23.
- ESPELAND Wendy Nelson et Berit Irene VANNEBO (2007), « Accountability, quantification, and law », *Annual Review of Law and Social Science*, 3(1), p. 21-43.
- FANG Xue (2022), « What contribution can the construction of the social credit system make to the dual carbon goal », *Social Credit System*, République populaire de Chine, Août, <https://credit.sc.gov.cn/xysc/index.shtml>, page consultée le 23 novembre 2025.
- FANG Xue (2023), « The ecological credit system is a new starting point for the harmonious coexistence of “man and nature” », *Baijiahao*, décembre, <https://baijiahao.baidu.com/s?id=1784515371312700887&wfr=spider&for=pc>, page consultée le 23 novembre 2025.
- FANG Xue, Gao GUILIN, Hou JIARU, Tan BOPING, Chen HAISONG, Liu MINGMING ... Chen WEIXIAN (2023), *The Construction of the Ecological Credit System by Many Experts*, Sichuan, Credit China Sichuan.
- FENG Hu (2017), « Rivers Flow in Me: Reflections From Zhejiang », *CWR*, 17 octobre, <https://www.chinawaterrisk.org/opinions/rivers-flow-in-me-reflections-from-zhejiang/>, page consultée le 23 novembre 2025.
- FOUCAULT Michel (2007/1977), *Security, Territory, Population: Lectures at the Collège de France, 1977-78*, New York, Springer.

- FOUCAULT Michel (2025/1972), *La société punitive. Cours au Collège de France (1972-1973)*, Paris, Seuil.
- FOURCADE Marion (2021), « Ordinal citizenship », *British Journal of Sociology*, 72(2), p. 152-173.
- FOURCADE Marion et Kieran HEALY (2013), « Classification situations: Life-chances in the neoliberal era », *Accounting, Organizations and Society*, 38(8), p. 559-572.
- FU Diana et Greg DISTELHORST (2018), « Grassroots Participation and Repression under Hu Jintao and Xi Jinping », *The China Journal*, 79, p. 100-122, DOI : <https://doi.org/10.1086/694299>.
- GAFFRIC Gwennaél et Jean-Yves HEURTEBISE (2013), « L'écologie, Confucius et la démocratie », *Écologie et politique*, 47(2), p. 51-61.
- GAO Guilin (2024), « Research on the credit supervision system of the whole chain and the whole life cycle », *The Construction of the Ecological Credit System, Interpreted by Many Experts*, <https://www.163.com/dy/article/JJ0CKGD5053257CG.html>, page consultée le 23 novembre 2025.
- GEALL Sean (2017), « Access to Environmental Information in China: Transparency, participation and the role of media », *Routledge Handbook of Environmental Policy in China*, Londres, Routledge, p. 248-262.
- GEALL Sam et Adrian ELY (2018), « Narratives and pathways towards an ecological civilization in contemporary China », *The China Quarterly*, 236(décembre), p. 1175-1196.
- GILLEY Bruce (2012), « Authoritarian environmentalism and China's response to climate change », *Environmental Politics*, 21(2), p. 287-307.
- GORON Coraline (2018), « Civilisation écologique et limites politiques du concept chinois de développement durable », *Perspectives chinoises*, 4, p. 41-55.
- GORON Coraline et Gillian BOLSOVER (2020), « Engagement or control? The impact of the Chinese environmental protection bureaus' burgeoning online presence in local environmental governance », *Journal of Environmental Planning and Management*, 63(1), p. 87-108, DOI : <https://doi.org/10.1080/09640568.2019.1628716>.
- GOURDON Julien, Stéphanie MONJON et Sandra PONCET, « Trade policy and industrial policy in China: What motivates public authorities to apply restrictions on exports? », *China Economic Review*, 40(septembre), DOI : <https://doi.org/10.1016/j.chieco.2016.06.004>.
- GRAMSCI Antonio (1971/1929-1935), *Selections from the Prison Notebooks*, Londres, Lawrence & Wishart.
- GREENPEACE (2014), « Un gâchis monstrueux : la pollution toxique de l'eau en Chine », *Greenpeace*, <https://www.greenpeace.org/international/story/6846/a-monstrous-mess-toxic-water-pollution-in-china/>, page consultée le 23 novembre 2025.
- HACKING Ian (1990), *The Taming of Chance*, Cambridge, Cambridge University Press.
- HAMMER Ben (2017), « Confucian thought and China's moral crisis », *International Communication of Chinese Culture*, 4(4), https://www.researchgate.net/publication/329117529_Confucian_Thought_and_China%27s_Moral_Crisis, page consultée le 7 mars 2023.
- HANSEN Mette Halskov, Hongtao LI et Rune SVARVERUD (2018), « Ecological civilization: Interpreting the Chinese past, projecting the global future », *Global Environmental Change*, 53, p. 195-203.
- HARDIN Garret, (1968), *Tragedy of the Commons*.
- HARVEY Tom (2021), *Xi Jinping, Carl Schmitt and China's New Era*, Oxford, Oxford House Research.

HAYEK Friedrich (1960), *The Constitution of Liberty*.

HEINZEKEHR Justin (2020), « Ecological civilization: The Political rhetoric of “Marxism with Chinese Characteristics” », *Non-Human Nature in World Politics: Theory and Practice*, 27 août, p. 163-177.

HEURTEBISE Jean-Yves (2023), « Green orientalism, brown occidentalism and chinese ecological civilization: Deconstructing the culturalization of the anthropocene to nurture transcultural environmentalism », *Asian Studies*, 11(2), p. 119-148.

HUANG Di et Xu LUO (2024), « Accélérer la construction d’une civilisation écologique numérique verte et intelligente », Résultat intermédiaire du projet Jeunes Chercheurs (n°21CKS034) *Recherche sur les caractéristiques de la « nature populaire » de la pensée de Xi Jinping sur la civilisation écologique et son mécanisme de mise en œuvre*, Fondation Nationale des Sciences Sociales de Chine, <http://www.xinhuanet.com/politics/20230703/812f33fcded5485a83d471e3b1d741da/c.html>, page consultée le 23 novembre 2025.

HUANG Jingyang et Kellee S. TSAI (2022), « Securing authoritarian capitalism in the digital age: The Political economy of surveillance in China », *The China Journal*, 88, DOI : <https://doi.org/10.1086/720144>.

HUANG Yanzhong (2020), *Toxic Politics: China’s Environmental Health Crisis and its Challenge to the Chinese State*, Cambridge, Cambridge University Press.

HUANG Yanzhong (2021), « Environmental Challenges of China’s Recovery After COVID-19 », *Time*, 2 february, <https://time.com/5935138/chinas-environment-economic-recovery/>, page consultée le 8 mars 2023.

JOHNS Fleur (2021), « Governance by data », *Annual Review of Law and Social Science*, 17, p. 53-71.

KER Michelle et Kate LOGAN (2014), « New environmental law targets China’s local officials », *China Dialogue*, 28 avril, <https://chinadialogue.net/en/pollution/6939-new-environmental-law-targets-china-s-local-officials/>, page consultée le 7 mars 2023.

KOSTKA Genia, Xuehua ZHANG et Kyoung SHIN (2019), « Information, technology, and digitalization in China’s environmental governance », 63(1), p. 1-13, DOI : <https://doi.org/10.1080/09640568.2019.1681386>.

LASCOUMES Pierre et Patrick LE GALÈS (2018), « Institutions, normes et instruments de l’action publique », dans *Sociologie de l’action publique*, Paris, Armand Colin, p. 84-102.

LI Vic et Graeme LANG (2010), « China's “green GDP” experiment and the struggle for ecological modernisation », *Journal of Contemporary Asia*, 40(1), p. 44-62.

LI Yifei et Judith SHAPIRO (2020), *China Goes Green: Coercive Environmentalism for a Troubled Planet*, Cambridge, Polity Press.

LISHUI CITY (2020), Lishui Ecological Credit Behavior Positive and Negative List, Trial ZJKC02-2020-0003, https://zjjcmspublic.oss-cn-hangzhou-zwynet-d01-a.internet.cloud.zj.gov.cn/jcms_files/jcms1/web3657/site/attach/0/fa370ef441f4476ba5c67d4db10300f4.pdf, page consultée le 23 novembre 2025.

LO Kevin (2020), « Ecological civilization, authoritarian environmentalism, and the eco-politics of extractive governance in China », *The Extractive Industries and Society*, 7(3), p. 1029-1035.

LOVEMAN Mara (2005), « The modern state and the primitive accumulation of symbolic power », *American Journal of Sociology*, 110(6), p. 1651-1683.

- MARQUIS Christipher, Jianjun ZHANG et Yanhua ZHOU (2011), « Regulatory uncertainty and corporate responses to environmental protection in China », *California Management Review*, 54(1), p. 39-63.
- MARINELLI Maurizio (2018), « How to build a “Beautiful China” in the anthropocene. The Political discourse and the intellectual debate on ecological civilization », *Journal of Chinese Political Science*, 23(3), p. 365-386.
- MINISTÈRE CHINOIS DES AFFAIRES ÉTRANGÈRE (2021), Les évolutions sur les émissions de CO₂ en Chine, Chine, https://www.fmprc.gov.cn/mfa_eng/topics_665678/2030kcxzfzyc/202109/P020211019152754484797.pdf, page consultée le 23 novembre 2025.
- MINISTÈRE DE L'ÉCOLOGIE ET DE L'ENVIRONNEMENT (2014), Cadre national sur l'évaluation du crédit environnemental des entreprises, République populaire de Chine, <https://www.mee.gov.cn/gkml/hbb/bwj/201401/W020140102498719710099.pdf>, page consultée le 23 novembre 2025.
- MINISTÈRE DE L'ÉCOLOGIE ET DE L'ENVIRONNEMENT (2014b), Corporate Environmental Credit Evaluation Measures, République Populaire de Chine, <https://www.mee.gov.cn/gkml/hbb/bwj/201401/W020140102498719710099.pdf>, page consultée le 23 novembre 2025.
- MONJON Stéphanie et Sandra PONCET (2018), *La transition écologique en Chine. Mirage ou « virage vert » ?*, Paris, Éditions Rue d'Ulm.
- MONJON Stéphanie et Élodie RENÉ (2024), « La civilisation écologique en Chine, entre outil de contrôle domestique et projet de rayonnement mondial », *Revue Internationale et Stratégique*, 135, p. 203-215.
- MONJON Stéphanie et Léa BOUDINET (2021), « The state of China's environment: What has changed in the past few years? », *Groupe d'études géopolitique*, septmebre, p. 128-133, <https://geopolitique.eu/en/articles/the-state-of-chinas-environment-what-has-changed-in-the-past-few-years/>, page consultée le 8 mars 2023.
- MONJON Stéphanie and Élodie RENÉ (2021), « The new tools of environmental governance in China: Top down control and environmental credit », *Green - Groupe d'Étude Géopolitique*, <https://geopolitique.eu/en/articles/the-new-tools-of-environmental-governance-in-china-top-down-control-and-environmental-credit/>, page consultée le 23 novembre 2025.
- MOUTON Gauthier (2022), « L'“Airpocalypse” en Chine : manifestations de la crise climatique et de la dépendance énergétique au charbon », *À propos*, 2, p. 68.
- OFFICE OF THE CENTRAL CYBERSPACE COMMISSION (2023), « Building a green and smart digital ecological civilization (commentator's observation) », *Accelerating the Construction of Digital China*, 6(avril), https://www.cac.gov.cn/2023-04/17/c_1683374705642157.htm, page consultée le 23 novembre 2025.
- RENÉ Élodie (2019), « La civilisation écologique contrôlée par le numérique en Chine », *Multitudes*, 76(3), p. 86-93
- ROSE Nikolas (1990), *Governing the Soul: The Shaping of the Private Self*, Londres, Taylor & Frances/Routledge.
- SCHMITT Carl (1921), *Die Diktatur. Von den Anfängen des modernen Souveränitätsgedankens bis zum proletarischen Klassenkampf* (La dictature. Des débuts de la pensée moderne de la souveraineté jusqu'à la lutte de classe prolétarienne), Berlin, Duncker & Humblot.
- SCOTT James C. (1998), *Seeing Like a State*, New Haven, Yale University Press, <https://yalebooks.yale.edu/book/9780300078152/seeing-like-a-state/>, page consultée le 5 mars 2023.

SEL Pierre (2019), « Chine : Comprendre le système de crédit social », *Revue Esprit*, octobre, <https://esprit.presse.fr/article/pierre-sel/chine-comprendre-le-systeme-de-credit-social-42317>, page consultée le 7 mars 2023.

SHI Zhiyuan et Chi JING (2024), « Promouvoir la modernisation chinoise par une civilisation écologique numérique », *Baidu*, <https://baijiahao.baidu.com/s?id=1788838250498613476&wfr=spider&for=pc>, page consultée le 20 novembre 2025.

STAMM Christopher (2021), « Les standards RSE dans une approche néo-institutionnelle », Working Paper, Montréal, Université de Montréal, DOI : 10.13140/RG.2.2.19443.22560.

SULLIVAN Lawrence R. (2023), « Tackling the environment: Xi Jinping and the pursuit of a “beautiful China” », *Journal of Environmental Assessment Policy and Management*, 25(1), DOI : <https://doi.org/10.1142/S1464333223500059>.

SUM Ngai-Ling (2019), « Ordoliberal authoritarian governance in China since 1978: World market, performance legitimacy, and biosovereign ordering », *South Atlantic Quarterly*, 118(2), p. 381-400.

SUN Yu et HUANG Vincent Guangsheng (2021), « Embedded data activism: the institutionalization of a grassroots environmental data initiative in China », *Chinese Journal of Communication*, 15(1), p. 115-137, DOI : <https://doi.org/10.1080/17544750.2021.1963997>.

SUPIOT Alain (2015), *La gouvernance par les nombres*, Paris, Fayard.

UNION EUROPÉENNE (2016), Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32016R0679>, page consultée le 23 novembre 2025.

UNITED NATIONS (2019), « Chinese initiative ant forest wins un champions of the earth award », United Nations Environment Program, <https://www.unep.org/news-and-stories/press-release/chinese-initiative-ant-forest-wins-un-champions-earth-award>, page consultée le 23 novembre 2025.

WANG Alex L. et al. (2018), « Explaining environmental information disclosure in China », *Ecology Law Quarterly*, 44(4), DOI : <https://doi.org/10.15779/Z386688J63>.

WANG Evelyn Li (2023), « A beacon of hope for a troubled Earth? A critical analysis of China's authoritarian approach to constructing “ecological civilization” », *Asia Pacific Journal of Environmental Law*, 26(1), p. 108-131.

WANG Ran (2019), « Experience and inspirations from the building of a beautiful Zhejiang », dans *Chinese Dream and Practice in Zhejiang - Ecology*, New York, Springer, p. 265-289.

WANG Xiaorui (2019), « The national ecological accounting and auditing scheme as an instrument of institutional reform in China: A discourse analysis », *Journal of Business Ethics*, 154(3), p. 587-603.

WANG Zhihe, Zanmei CUI, Wenxi ZHANG et Qing TANG (2022), « Organic Marxism in China: The challenges and their origins », *Balkan Journal of Philosophy*, 14(1), p. 23-34, DOI : <https://doi.org/10.5840/BJP20221414>.

WESTMAN Linda et Ping HUANG (2022), « Ecological civilisation: A historical perspective on environmental policy narratives in China », *International Quarterly for Asian Studies (IQAS)*, 53(2), p. 181-206.

WU Karen (2023), « A brief analysis of the application of ecological credit in financial assistance in realizing the value of ecological products », *Central University of Finance and Economics-Bank of*

Beijing Dual Carbon and Finance Research Center, <https://m.163.com/dy/article/HSSH98DU05199DKK.html>, page consultée le 23 novembre 2025.

XINHUA (2023), « Growing public participation in “carbon inclusion” empowers China's green development », *China Daily*, 18 juillet, <https://global.chinadaily.com.cn/a/202307/18/WS64b65138a31035260b8171a0.html>, page consultée le 23 novembre 2025.

YU Ivy (2022), « Alibaba Launches Carbon Ledger to Drive Eco-friendly Consumer Behavior », *Alizila*, 8 août, <https://www.alizila.com/alibaba-launches-carbon-ledger-to-drive-eco-friendly-consumer-behavior/>, page consultée le 23 novembre 2025.

YU Sun et Wenjie YAN (2020), « The power of data from the global south: Environmental civic tech and data activism in China », *Journal of Communication*, 14, <https://ijoc.org/index.php/ijoc/article/view/13440>, page consultée le 8 mars 2023.

ZENG Jing, Chung-hong CHAN et King-wa FU (2017), « How social media construct “truth” around crisis events: Weibo’s rumor management strategies after the 2015 Tianjin blasts », *Policy & Internet*, 9(3), p. 297-320, DOI : <https://doi.org/10.1002/POI3.155>.

ZHANG Chenchen (2020), « Governing (through) trustworthiness: technologies of power and subjectification in China’s social credit system », *Critical Asian Studies*, 52(4), p. 565-588, DOI : <https://doi.org/10.1080/14672715.2020.1822194>.

ZHANG Lei, Arthur PJ MOL et Guizhen HE (2016), « Transparency and information disclosure in China’s environmental governance », *Current Opinion in Environmental Sustainability*, 18, p. 17-24, DOI : <https://doi.org/10.1016/J.COSUST.2015.03.009>.

ZHOU Di (2020), « China’s environmental vertical management reform: An effective and sustainable way forward or trouble in itself? », *Laws*, 9(4), p. 25, DOI : <https://doi.org/10.3390/LAWS9040025>.

NOTES

1. Xue Fang est secrétaire général adjoint de la Beijing Credit Society, chercheur au Centre de recherche sur la législation du crédit et l'évaluation du crédit de l'Université normale de la capitale.
2. En effet, si le référentiel expose dans le détail la méthode de construction de la notation, la répartition des points et le type de critères à prendre en compte pour chaque composante, il ne fixe pas le mode de calcul des indicateurs. L'intégration d'un critère peut donc autant ressortir d'une évaluation matérielle directe, que d'une méthode formelle à partir d'une mesure.
3. À l'occasion de cet article, nous avons effectué une évaluation plus récente sur la même plateforme (par sondage, en comptant le nombre d'occurrences de chaque niveau de classification), et les résultats sont cohérents avec ceux obtenus en 2021.

RÉSUMÉS

Cet article explore comment les ambitions du Parti Communiste Chinois (PCC) en matière de contrôle des données environnementales se manifestent à la fois sur le plan discursif, à travers les récits du Parti sur la Civilisation Écologique Numérique (CEN), et sur le plan opérationnel, via le Système de Crédit Écologique (SCE) qui vise à mesurer et noter le niveau « d'intégrité écologique » des entreprises et des individus afin d'encourager les comportements écoresponsables tout en sanctionnant ceux jugés nuisibles à l'environnement. Deux sous-systèmes du SCE ont été étudiés : le système de crédit environnemental des entreprises et le système de crédit écologique des citoyens. Cette analyse permet d'interroger les fondements idéologiques du SCE ainsi que son efficacité opérationnelle pour atteindre les objectifs climatiques fixés par Pékin : un pic des émissions d'ici 2030 et la neutralité carbone à l'horizon 2060. Les concepts clés mobilisés par les chercheurs chinois travaillant à la conceptualisation de la CEN et du SCE ont ainsi été étudiés au prisme des Critical Data Studies. Les textes retenus reflètent les productions académiques, idéologiques et doctrinales de chercheurs chinois et d'experts institutionnels occupant des positions clés dans les organismes de recherche ou de pilotage de la CEN et du SCE en Chine.

This article explores how the Chinese Communist Party's (CCP) ambitions to control government data are manifested at both the discursive level, through the Party's narratives on Ecological Civilization (EC), and at the operational level, through the Ecological Credit System (ECS), which aims to measure and rate the "ecological integrity" level of companies and individuals in order to encourage environmentally responsible behaviour while penalizing behaviour deemed harmful to the environment. Two subsystems of the ECS were studied: the corporate environmental credit system and the citizens' ecological credit system. This analysis makes it possible to question the ideological foundations of the ECS and its operational effectiveness in achieving the climate goals set by Beijing: peak emissions by 2030 and carbon neutrality by 2060. The key concepts used by Chinese researchers working on the conceptualization of the EC and the ECS were thus studied through the lens of Critical Data Studies (CDS).

Este artículo examina cómo, las ambiciones del Partido Comunista Chino (PCC), en lo concerniente al control de los datos gubernamentales, se reflejan por una parte, a nivel discursivo, por intermedio de las narraciones del Partido sobre la Civilización Ecológica (CEN) y por otra parte, a nivel operativo, por intermedio del Sistema de Crédito Ecológico (SCE) que tiene como objetivo tomar nota sobre el grado de "integridad ecológica" de las empresas y personas con el propósito de motivar comportamientos ecoresponsables y sancionar a las que son perjudiciales al medio ambiente. Para realizar el análisis, se estudiaron dos subsistemas del SCE, es decir, el sistema de crédito para el medio ambiente en las empresas y el sistema de crédito para la ecología humana. Este análisis permite examinar los fundamentos ideológicos del SCE y la eficacia operativa para lograr los objetivos climáticos establecidos por Pékin, lo que significa, un punto máximo de las emisiones desde ahora hasta 2030 y la neutralidad de carbono para el año 2060. Los conceptos clave utilizados por los investigadores chinos que trabajan en la conceptualización de CEN y del SCE fueron estudiados a través del prisma de *Critical Data Studies* (CDS).

China, Civilización Ecológica Digital, datos medioambientales, integridad ecológica, Sistema de Crédito Ecológico

INDEX

Keywords : China, Digital Ecological Civilization, environmental data, ecological integrity, Ecological Credit System

Mots-clés : données environnementales, Chine, Civilisation Écologique Numérique, Système de Crédit Écologique, intégrité écologique

AUTEURS

ÉLODIE RENÉ

Elodie René est docteure en sciences politiques (l'Université de Northampton), spécialiste de la gouvernance environnementale et des systèmes de crédit social et écologique en Chine. Elle est chercheuse au Netlab. Courriel : elodie.rene@ananke.coop

SAMUEL TRONÇON

Samuel Tronçon est docteur en philosophie, spécialiste en informatique appliquée aux sciences sociales. Il est chercheur au Netlab, après avoir été affilié à l'Université Paris 8 et à l'Institut de Mathématiques de Marseille. Courriel : stroncon@ananke.coop

Pratiques numériques et mise en données de la défense des droits humains

Vers une datafication de nos affects ?

Normand Landry, Julien Hocine, Simon Fitzbay et Marie-Josée Dupuis

- 1 Une abondante littérature témoigne de l'intégration progressive des technologies numériques dans les activités des défenseur·es des droits humains. Ces technologies sont utilisées pour enregistrer et diffuser des preuves de violations de droits fondamentaux, communiquer des informations de manière confidentielle, s'adresser à différents publics et mobiliser des alliés (Ziccardi, 2012 ; McPherson, 2015 ; Meikle, 2018). Ainsi les technologies numériques s'insèrent-elles dans les pratiques constitutives de la défense des droits humains et dans les processus d'organisation des mouvements de luttes sociopolitiques qui les caractérisent (Kasm, 2018 ; Kazansky, 2015 ; Michaelsen, 2018).
- 2 Ces mêmes technologies sont néanmoins associées à d'importants risques pour les personnes qui militent en faveur des droits humains. Leurs adversaires les emploient couramment à des fins d'intimidation, d'infiltration de leurs réseaux et de désinformation (Chiweshe et Hove, 2023 ; Gregory, 2019). Les États les utilisent également pour identifier des individus, intercepter des communications et porter atteinte à la vie privée ainsi qu'à la sécurité de celles et ceux qui œuvrent pour la protection des droits fondamentaux (Anstis, 2024). Les agences internationales de défense des droits humains et les organisations à but non lucratif actives dans ce domaine signalent des menaces persistantes et croissantes liées à la répression numérique des personnes engagées dans cette lutte (Markland, 2021).
- 3 Les technologies numériques constituent assurément des instruments de lutte dans le cadre de la défense des droits humains — servant tant à l'expansion des pratiques militantes qu'à l'application de mesures répressives — et des enjeux de lutte en elles-mêmes. À ce titre, les défenseur·es des droits humains cherchent à orienter les cadres

réglementaires qui encadrent les réseaux de communication numérique, mais aussi à accroître et à sécuriser leur accessibilité (Joyce, 2015 ; Khamdamova, 2020).

- 4 Dans le contexte de la numérisation accélérée de la défense des droits humains, les processus de mise en données jouent un rôle déterminant. Ils contribuent à redéfinir les répertoires d'action des militant·e·s et des organisations engagées, tout comme ceux des acteurs impliqués dans leur répression. Ces processus montrent des moyens novateurs de collecte, de traitement et d'usage des informations mobilisées par les défenseur·e·s des droits humains pour créer des récits, reconstituer des scènes et étayer des preuves de violations des droits humains (Gregory, 2019 ; Gutierrez, 2021). Toutefois, ces processus mettent en lumière un ensemble de vulnérabilités exploitables par leurs adversaires que cet article recensera. Inscrite dans la sociologie des usages des technologies de l'information et de la communication (Denouël et Granjon, 2011 ; Kane, 2013), notre approche permet de recenser, d'une part, les usages délibérés, récurrents, orientés des technologies numériques à des fins de défense des droits humains et, d'autre part, la répression par des acteurs étatiques des regroupements et individus investis dans des activités de défense des droits humains.
- 5 Cet article présente les résultats d'une revue systématique de la littérature sur plus de dix ans (2010-2021) portant sur les pratiques numériques de défense et de répression des droits humains. L'analyse de la littérature met en évidence des répertoires de pratiques numériques à la fois dans la défense des droits humains et dans la répression des personnes et des organismes qui s'y consacrent. Elle permet par ailleurs de dresser un inventaire des facteurs qui exercent une influence sur la capacité des défenseur·e·s des droits humains à mobiliser les technologies numériques dans le cadre de leurs activités. Les constats indiquent que les processus de mise en données sont déterminants dans les pratiques des défenseur·e·s des droits humains et des États impliqués dans leur répression. Ces processus jouent également un rôle crucial dans la détermination des possibilités et des risques des technologiques en contexte de militance démocratique. L'analyse de la littérature met en évidence des représentations et des actions politiques entreprises par les défenseur·e·s des droits humains pour exercer une pression sur les processus réglementaires et de gouvernance des données numériques dans le but de réduire les risques et d'accroître leur capacité d'action¹.
- 6 Cet article est divisé en quatre temps. Il introduira les principales assises conceptuelles et théoriques de l'étude avant de la présenter en détail. Il présentera ensuite la méthode de recherche et les résultats de l'étude. Enfin, une discussion abordera les questions et les enjeux associés à la mise en données de la défense des droits humains.

Assises conceptuelles et théoriques

- 7 Nos analyses s'appuient sur l'articulation de quatre concepts clés : les défenseur·e·s des droits humains, les pratiques, les répertoires et la mise en données. Cette approche conceptuelle permet d'éclairer les processus de numérisation prenant place au sein des pratiques stabilisées de défense et de répression des droits humains.

Défenseur·e·s des droits humains

- 8 L'expression « défenseur·e·s des droits humains² » est traitée de manière volontairement vague à l'échelle internationale. Elle désigne toute personne qui,

individuellement ou avec d'autres, agit à des fins de protection et de réalisation des droits humains, sur le plan tant national qu'international (Nations Unies, 1999, art. 12). Ainsi, les individus et regroupements militants actifs sur une pluralité d'enjeux — allant de l'égalité des genres à la lutte environnementale — peuvent revendiquer ou se voir attribuer cette appartenance s'ils cadrent, ou si l'on cadre pour eux, la nature de leurs activités en termes de défense de droits humains. Si l'absence de définition claire et balisée perpétue une ambiguïté et des controverses sur les spécificités et les frontières du champ que l'expression couvre (Kogan, 2013 ; Nah et al., 2013), elle évite également d'exclure d'une reconnaissance officielle, d'une légitimation et, surtout, de mesures de protection les individus et regroupements actifs par rapport à une panoplie d'enjeux qui touchent aux droits humains.

- 9 Le Rapporteur spécial sur la situation des défenseur·es des droits humains précise un ensemble de caractéristiques essentielles à la pratique. Il souligne la reconnaissance du caractère universel des droits humains et la nécessité d'adhérer à l'intégralité des droits. L'action des défenseur·es s'inscrit à plusieurs échelons, du local à l'international, et comprend diverses activités : la collecte et la diffusion d'informations sur les violations des droits humains, l'assistance aux victimes, la mise en cause des responsables de ces violations, la participation aux réformes de gouvernance, l'application des instruments internationaux de protection et l'éducation aux droits humains (Nah et al., *op. cit.*). Les actions militantes menées par les défenseur·es des droits humains doivent être non violentes. La littérature scientifique et les institutions internationales insistent sur la nécessité de reconnaître tant les formes « professionnalisées » de défense des droits humains — caractérisées par une expertise, des emplois et des ressources spécifiques — que les engagements plus informels et ponctuels, souvent bénévoles, qu'assument des individus ou des collectifs concentrés sur des enjeux ciblés, et faiblement structurés (Kogan, *op. cit.* ; Nah et al., *op. cit.* ; Van der Vet et Lyytikäinen, 2015 ; Land, 2009). Enfin, la question de la gestion du risque encadre les activités de défense des droits humains, les individus impliqués étant régulièrement exposés à l'incarcération, au harcèlement, à la diffamation, ainsi qu'à des violences physiques et psychologiques (Al-Rawi, Tenove et Klein, 2023 ; Gregory, 2019 ; Gutierrez, *op. cit.* ; Van der Vet et Lyytikäinen, *op. cit.*).

Pratiques

- 10 De manière générale, le concept de pratiques renvoie à des activités organisées récurrentes ou routinisées permettant d'atteindre des objectifs identifiés par les individus qui y ont recours. Une pratique s'actualise dans des manières de faire, des processus, des normes qui encadrent, balisent et définissent des activités : elle s'inscrit dans les logiques sociales et les systèmes de significations de ceux et celles qui la portent (Certeau, 1984). Le champ couvert par l'expression « défense des droits humains » peut se concevoir sous l'angle de quatre catégories de pratiques qui en sont constitutives.
- 11 En premier lieu, les activités par lesquelles les informations sur les violations de droits humains sont recueillies, collectées, regroupées, mises en relation et organisées de manière à constituer des preuves, témoignages ou supports argumentatifs, sont essentielles à la défense des droits humains. Ces activités se regroupent en des pratiques d'enquête et de consignation basées sur des normes éthiques et juridiques,

des manières de faire et la disponibilité d'outils et de ressources (Fletcher, 2022). Elles sont essentielles pour rendre imputables les responsables des violations de droits humains, soit devant l'opinion publique ou les tribunaux.

- 12 En deuxième lieu, les activités communicationnelles, que nous regroupons sous le vocable de pratiques de « plaider », constituent une dimension centrale de la défense des droits humains. La notion de plaider renvoie ici à l'élaboration de discours et de récits de manière à convaincre, à faire la démonstration de situations où des violations de droits humains sont constatées et exigent correction ou réparation (Land et Aronson, 2020 ; Fletcher, *op. cit.*). Les publics usuels de ces plaidoyers se retrouvent dans des milieux juridiques, au sein d'élites sociales, économiques, médiatiques ou politiques de la société civile, ou sont constitués des populations dont les représentations sont constitutives d'une opinion publique.
- 13 En troisième lieu, les défenseur·es des droits humains s'appuient principalement sur des modes d'action éprouvés pour exercer une pression sur les autorités publiques — tels que la manifestation, la pétition ou la mise en scène d'événements spectaculaires visant à susciter une réaction politique ou médiatique. Les formes de dissidence plus disruptives, comme le sabotage, la divulgation d'informations choquantes ou confidentielles ou encore encore les attaques informatiques, se situent généralement en dehors des répertoires d'action jugés acceptables et traditionnellement mobilisés dans la défense des droits humains, bien que certains collectifs s'en réclament ponctuellement³ (Sorell, 2015 ; Kingston et Stam, 2013 ; Chak, 2014). Ces pratiques contentieuses visent à établir un certain rapport de force, une capacité à contraindre des acteurs publics et à entreprendre des actions correctives ou remédiatrices ou encore à faire cesser des actions jugées illégitimes ou inacceptables.
- 14 En quatrième et dernier lieu, les défenseur·es des droits humains s'investissent dans des activités de protection de personnes ou de groupes soumis à des traitements arbitraires et oppressifs, des violences, des injustices ou des persécutions. Ces activités comprennent la défense en justice de personnes ou de groupes criminalisés, la mise en visibilité publique de victimes, l'établissement de réseaux de soutien pour ces dernières, leur dissimulation ou leur déplacement à l'extérieur d'une juridiction qui leur est hostile (Stefanovska, 2024 ; Makhortykh, 2023).
- 15 Ces grandes catégories de pratiques n'épuisent pas le spectre des activités accomplies par les défenseur·es des droits humains. Elles permettent toutefois de définir les contours du champ d'action associé au statut de « défenseur·es des droits humains ».

Répertoires

- 16 La sociologie des mouvements sociaux mobilise le concept de répertoires d'action collective (*repertoires of contention* ; Tilly, 1995 ; Tilly et Tarrow, 2007) pour illustrer des « routines acquises et partagées de stratégies et tactiques politiques concevables et disponibles pour des activistes à un moment et à un lieu donnés » (Landry, Aubin et Sénécal, 2015, p. 157). Les répertoires d'action collective évoluent selon les milieux et le temps, les ressources disponibles, les contraintes externes et structurelles ainsi que la capacité des individus et des groupes à se saisir des occasions s'offrant à eux.
- 17 Le développement fulgurant des technologies numériques au cours des deux dernières décennies est venu reconfigurer et élargir les répertoires d'action collective. Une littérature abondante aborde les processus d'appropriation militante des

fonctionnalités des dispositifs numériques⁴ (Granjon, 2017 ; Milan, 2013). Ces mêmes dispositifs, et les infrastructures numériques qui en assurent le fonctionnement, sont néanmoins instrumentalisés par des États à des fins d'identification, de surveillance, d'intimidation, de harcèlement et d'oppression des individus et des groupes avec lesquels ils entretiennent des rapports conflictuels (Feldstein, 2021). Nous mobilisons ainsi le concept de répertoires de pratiques numériques pour illustrer des usages intentionnels et récurrents de technologies numériques à des fins politiques. Les objectifs peuvent être, selon les contextes, orientés autour de la défense des droits humains ou de la répression des personnes qui s'y consacrent.

Mise en données

- 18 Le concept de mise en données évoque la somme des processus par lesquels des activités humaines sont transformées en information numérique pouvant être exploitée (Casemajor, 2023). Une mise en données implique l'extraction, la collecte ou la production du « matériel brut » qui sera transformé, un processus de conversion de ce matériel en format numérique, un traitement numérique de ce matériel pour le rendre pertinent et significatif au vu d'objectifs spécifiques ainsi que sa valorisation dans le cadre d'activités précises (Dijck, 2014). Si le concept de mise en données a d'abord été associé à une transformation des aspects de la vie quotidienne en informations numériques (Office québécois de la langue française, 2014), il est désormais acquis qu'il n'existe pratiquement plus de secteurs de l'activité humaine qui y échappent.
- 19 Les processus et effets d'une mise en données des sociétés sont profondément politiques. S'ils sont susceptibles de favoriser des formes renouvelées de participation citoyenne et un élargissement des répertoires d'action collective (Bhargava et al., 2015 ; Milan et Van der Velden, 2016), la littérature témoigne de l'intégration des données dans des rapports de domination médiatisés par le numérique (Zuboff, 2019). Dans le cadre de notre étude, nous mobilisons le concept de mise en données pour analyser la manière dont la littérature savante révèle des processus par lesquels les défenseur·es des droits humains extraient, collectent, produisent, mettent en relation et diffusent des informations numériques dans leurs pratiques. Cela nous permet d'examiner la place des données dans les répertoires de pratiques numériques de défense des droits humains, puis de contextualiser et de problématiser cette intégration.

Méthode

- 20 L'analyse systématique de la littérature de langue anglaise couvre une décennie cruciale dans la transformation des pratiques numériques des défenseur·es des droits humains et de leurs adversaires (2010-2021). Cette décennie a vu s'intégrer de manière accélérée les technologies numériques et les militances démocratiques à grande échelle (Costanza-Chock, 2020 ; Lievrouw, 2023 ; Milan, *op. cit.* ; Meikle, *op. cit.* ; McNutt, 2018) tout en mettant en lumière des dynamiques répressives et de contrôle soutenues par des technologies et infrastructures numériques de la part des États et de partenaires privés (Morozov, 2011 ; Zuboff, 2019 ; Rudie et al., 2021 ; Lyon, 2015).
- 21 Huit bases de données ont été explorées pour conduire cette analyse⁵. Une sélection de mots-clés variés a été retenue en tenant compte de la pluralité des expressions utilisées

pour évoquer le champ d'activité et les technologies couverts⁶. Pour accroître l'efficacité du processus d'identification des articles pertinents, six revues apparaissant de manière récurrente dans les résultats de recherche par mots-clés ont été examinées de manière systématique⁷. Les bibliographies de l'ensemble des articles ayant été classés comme pertinents pour l'étude ont été recensées systématiquement pour enrichir le corpus primaire. Un processus rigoureux à double arbitrage a été effectué dans le but de circonscrire un corpus primaire, composé d'articles directement liés à notre étude (n = 51). Nous avons également assemblé un corpus secondaire (n = 79) constitué d'articles présentant des informations pertinentes du point de vue des enjeux et du contexte plus large des droits humains, mais où les pratiques numériques des défenseur·es des droits humains sont absentes ou effleurées.

- 22 Le corpus primaire a été traité suivant une pratique bien établie de codage itératif à double arbitrage (Locke, Feldman et Golden-Biddle, 2022). Celle-ci a conduit à la constitution d'une grille d'analyse constituée des volets suivants :
 1. Les catégories de discours universitaires à l'égard de la contribution des technologies numériques à la défense des droits humains ;
 2. Les répertoires de pratiques mobilisant les technologies numériques pour la défense des droits humains ;
 3. Les répertoires de pratiques répressives soutenues par les technologies numériques ;
 4. Les catégories de facteurs intervenant dans le potentiel d'usage des technologies numériques ;
 5. Les catégories de transformations de la pratique de la défense des droits humains jugées liées à la numérisation du champ.
- 23 Le corpus a été codé intégralement suivant les standards de rigueur et de fidélité interjuges qui s'imposent en analyse qualitative (Cascio et al., 2019 ; Paul, Bezerra Rodrigues et Seniuk Cicek, 2021 ; Giesen et Roeser, 2020). Le travail de codage effectué a conduit à la constitution de regroupements hiérarchisés de catégories pour tous les volets susmentionnés. Ces catégories varient dans le nombre et la diversité de leurs éléments constitutifs ; elles témoignent à la fois du spectre des thèmes associés à chaque élément de la grille d'analyse et de leur « poids » au sein la littérature analysée. Ce poids relatif est établi par le nombre d'articles distincts en faisant mention.
- 24 Dans le cadre du présent article, nous présenterons des résultats qui mettent en lumière des répertoires de pratiques mobilisant les technologies numériques pour la défense des droits humains, des répertoires de pratiques répressives soutenues par les technologies numériques ainsi que des catégories de facteurs intervenant dans le potentiel d'usage des technologies numériques. La présentation de ces résultats permettra de tenir une discussion plus large sur les effets des processus de mise en données dans les activités de défense des droits humains.

Résultats

- 25 L'analyse de la littérature permet le regroupement de pratiques numériques en catégories distinctes se divisant selon qu'elles sont portées *par* ou *contre* des défenseur·es des droits humains. Les regroupements de pratiques soulevés par la littérature savante permettent de circonscrire quatre grands répertoires de pratiques numériques mobilisées par les défenseur·es des droits humains et trois grands

répertoires de pratiques numériques répressives mobilisées par les États. Chaque grande catégorie regroupe un ensemble de catégories de pratiques plus fines (voir tableau 1).

Tableau 1. Répertoires de pratiques numériques

Répertoires de pratiques numériques de défense des droits humains	Nombre de pratiques distinctes recensées
Pratiques communicationnelles et informationnelles (n = 30)	15
Pratiques organisationnelles (n = 24)	12
Pratiques d'enquête et de consignation (n = 21)	11
Pratiques contentieuses (n = 14)	8
Répertoires de pratiques numériques répressives	
Censure (n = 27)	7
Surveillance étatique (n = 23)	6
Propagande (n=18)	6

Légende : « n » renvoie ici à un nombre d'articles distincts tirés du corpus primaire.

26 Il est à noter que la recension des pratiques ne constitue pas un inventaire tiré d'une analyse de pratiques constatées ou déclarées, mais identifie des discours tenus par la communauté universitaire au sujet des activités numériques. Elle témoigne ainsi de l'attention portée par le milieu universitaire sur certaines pratiques et catégories de pratiques ; elle peut être interrogée quant à sa complétude, à son exactitude, à l'accent mis sur certaines pratiques ou à la faible considération accordée à d'autres. Nous y reviendrons en conclusion.

Répertoires de pratiques numériques de défense des droits humains

27 Les répertoires de pratiques numériques de défense des droits humains rassemblent des activités qui soutiennent des actions de défense des droits et des victimes ou visant l'avancement des causes sociales. Ils regroupent au total 46 pratiques distinctes classées dans l'une des 4 catégories. Le tableau 2 présente, pour chacune d'elles, les cinq pratiques les plus couramment mentionnées par la littérature.

Tableau 2. Principales pratiques numériques par catégories

Pratiques communicationnelles et informationnelles	Pratiques organisationnelles
--	------------------------------

• Mobiliser les citoyens pour fournir des preuves visuelles et accroître le pouvoir symbolique de la lutte (n = 12) • Développer des espaces numériques favorisant la communication entre les militants et leurs soutiens (n = 9) • Utiliser les technologies comme outils d'organisation et de plaidoyer (n = 7) • Développer une « logique de présence » pour être visible dans des environnements numériques (n = 6) • Diffuser des mèmes et des mots-clics pour rassembler les conversations sur les questions des droits humains (n = 5)	• Renforcer les capacités des défenseur·es des droits humains par des formations et des outils pratiques (n = 6) • Utiliser Internet pour maintenir les réseaux entre militants (n = 6) • Mettre en place des initiatives collectives pour renforcer les pratiques numériques en matière de défense des droits humains (n = 4) • Utiliser des communications chiffrées pour garantir la confidentialité des messages (n = 4) • Plaider pour influencer le développement des technologies (n = 3)
Pratiques d'enquête et de consignation	Pratiques contentieuses
• Conduire des analyses médico-légales multidimensionnelles pour la consignation des violations des droits humains (n = 14) • Utiliser les téléphones mobiles pour la consignation par témoins oculaires (n = 7) • Utiliser des informations en provenance de sources multiples et ouvertes à des fins d'enquête et de contextualisation (n = 7) • Utiliser des drones pour recueillir des preuves visuelles et mener des interventions humanitaires (n = 4) • Utiliser l'imagerie satellite pour répertorier les violations dans le monde entier (n = 3)	• Utiliser des tactiques de contre-surveillance lors de mise en visibilité d'abus (n = 7) • Mettre en place des tactiques de « hacktivisme » à des fins de changement social (n = 3) • Créer des contre-récits avec le data-activisme (n = 2) • Imputer une responsabilité aux plateformes ou aux fournisseurs d'accès Internet (n = 2) • Se réappropriier les infrastructures énergétiques pour une meilleure résilience en cas de coupures (n = 1)

- 28 Les pratiques communicationnelles et informationnelles sont à la fois les plus nombreuses et les plus couramment référencées par la littérature. Elles témoignent d'un accent mis sur des activités de consignation et de diffusion de violations de droits humains (Deutch et Al Khatib, 2018 ; Garcia *et al.*, 2020 ; Saber et Rahman al-Jaloud, 2024) ainsi que sur la création d'espaces numériques investis par les défenseur·es des droits humains pour alimenter la mobilisation et renforcer leurs réseaux (Brough et Li, 2013 ; Deutch et Al Khatib, *op. cit.* ; Michaelsen, *op. cit.*). Les plateformes de réseaux sociaux sont d'ailleurs utilisées pour communiquer avec des publics et remédier à des insuffisances et à des insatisfactions en matière de couverture médiatique (Garcia *et al.*, *op. cit.*). La littérature souligne également des tactiques basées sur le dévoilement d'informations confidentielles et des activités visant à « humilier » des adversaires en exposant publiquement des malversations, des abus et des injustices avérés (Dutt et Rasul, 2014 ; Markland, *op. cit.* ; Van der Vet et Lyytikäinen, *op. cit.*).
- 29 Les pratiques organisationnelles recensées par la littérature regroupent des activités numériques qui contribuent à l'organisation, à l'efficacité et à la sécurité des défenseur·es des droits humains. Elles révèlent des architectures organisationnelles et des réseaux numériques sur lesquels s'appuient les défenseur·es des droits humains dans le cadre de leurs activités (Deutch et Al Khatib, *op. cit.* ; Tenove, 2019). Ces

pratiques intègrent le développement de formations en ligne, le partage des connaissances et l'innovation en matière de défense des droits humains (Gregory, 2010 ; Sorell, *op. cit.*) ainsi que le développement d'outils cryptographiques garantissant la confidentialité des communications (Joyce, *op. cit.* ; Kazansky, *op. cit.* ; Khamdamova, *op. cit.*). Enfin, les défenseur·es des droits humains militent pour l'intégration de normes respectueuses des droits humains auprès d'acteurs réglementaires et de sociétés commerciales qui conçoivent et distribuent des services, des dispositifs et des applications numériques (Maréchal, 2015 ; Mueller et Badiei, 2019).

- 30 Les pratiques d'enquête et de consignation constituent des activités centrales à la défense de droits humains. La crédibilité des accusations et des dénonciations publiques effectuées par les défenseur·es des droits humains est tributaire de la rigueur et de la fiabilité de ces pratiques : les technologies numériques sont ainsi employées pour reconstituer des scènes de crime, consigner, développer et étayer des preuves (Deutch et Al Khatib, *op. cit.* ; Herscher, 2014), authentifier et procéder à l'analyse d'informations (Gregory, 2019 ; Van der Velden, 2015). L'usage de drones et de l'imagerie satellite permet d'accéder à des sites dont l'accès est restreint et d'y collecter des informations qui appuieront la constitution de rapports documentés (Freeman, 2021 ; Gutierrez, *op. cit.*).
- 31 Enfin, les pratiques contentieuses regroupent l'ensemble des activités numériques qui visent à établir des rapports de force avec des adversaires ou à mettre sous pression ces derniers. Le terme *pratiques contentieuses* est tiré de la sociologie des mouvements sociaux. Il réfère à des dynamiques conflictuelles par lesquelles s'expriment des revendications qui ciblent des gouvernements ou qui exigent que ces derniers entreprennent des mesures à l'égard de tierces parties (voir Tilly, 2008). Les pratiques contentieuses recensées par la littérature comprennent un ensemble de tactiques associées à des formes de « hacktivisme » (attaques sur des sites et des infrastructures numériques, intrusion et vol de données, détournement ou défigurement de sites Web) (Chak, *op. cit.* ; Kingston et Stam, *op. cit.* ; Sorell, *op. cit.*) ainsi que des injonctions à l'égard d'entreprises technologiques qui autorisent ou facilitent des violations de droits humains (Markland, *op. cit.*). Les études attestent aussi du rôle important des activités de contre-surveillance — de consignation et de dénonciation d'abus de toute sorte — dans la mise sous pression et en imputabilité publique d'adversaires gouvernementaux (Maréchal, *op. cit.*).

Répertoires de pratiques numériques répressives

- 32 Les répertoires de pratiques numériques répressives décrivent des mesures mobilisées par les autorités publiques pour prévenir, réprimer et décourager les activités associées à la défense des droits humains⁸. Ils regroupent au total 19 pratiques distinctes classées dans des catégories associées à la censure, à la propagande et à la surveillance étatique. Le tableau 3 présente, pour chacune d'elles, les cinq pratiques les plus couramment mentionnées par la littérature.

Tableau 3. Principales pratiques numériques par catégories

Pratiques de censure	Pratiques de propagande	Pratiques de surveillance étatique
----------------------	-------------------------	------------------------------------

• Censurer Internet pour contrôler et filtrer l'information en ligne (n = 12) • Procéder à des coupures tactiques d'Internet (n = 12) • Intimider les activistes en ligne (n = 9) • Adopter des lois sur la désobéissance civile électronique et restreindre l'utilisation des technologies numériques (n = 8) • Contraindre à l'autocensure (n = 6)	• Mettre en place des trolls commandités par l'État et créer de faux récits (n = 6) • Contrôler et utiliser les médias d'État pour diffuser la propagande en faveur du régime (n = 5) • Orchestrer des campagnes de diffamation pour intimider les activistes et les mettre en danger (n = 5) • Manipuler ou falsifier du contenu en ligne (n = 4) • Diaboliser l'usage des réseaux sociaux (n = 2)	• Surveillance et suivi des activités en ligne des dissidents (n = 12) • Pratiques diverses de piratage (n = 11) • Recueillir des informations privées sensibles (n = 10) • Utiliser des technologies de reconnaissance faciale pour identifier les dissidents (n = 6) • Développer des logiciels et des appareils pour intercepter les données des utilisateurs (n = 4)
--	---	--

- 33 Les coupures d'Internet ainsi que le contrôle et le filtrage des informations en ligne figurent parmi les pratiques de censure les plus fréquemment mentionnées dans notre corpus. Kaun et Tréré (2020) s'appuient sur la classification de Howard et al. (2011) pour distinguer plusieurs formes de restrictions appliquées aux réseaux numériques. Celles-ci vont de la restriction totale d'accès aux réseaux dans un territoire donné à l'inaccessibilité ciblée de sites Web spécifiques, tout en y pratiquant l'identification et l'exclusion d'utilisateurs spécifiques. D'autres mesures telles que la réduction draconienne de la vitesse de connexion entravent également l'accès aux services et ressources numériques, altérant leur efficacité et leur fonctionnement. Les fournisseurs d'accès à Internet sont fréquemment sollicités par les États, en tant qu'intermédiaires pour rendre leurs services inopérables (Siatitsa, 2020). Parallèlement, les tactiques d'intimidation contre les défenseur·es des droits humains se sont intensifiées avec l'essor des réseaux sociaux, y compris le harcèlement et la formulation de menaces de violences physiques (East and Horn of Africa Human Rights Defenders Project, 2013 ; Hankey et Ó Clunaigh, 2013 ; Onuch, 2015). En outre, les autorités exigent des entreprises privées l'accès aux données des usagers (Nyst, 2018 ; Siatitsa, *op. cit.*) et s'appuient sur des cadres juridiques criminalisant la dissidence en ligne par le biais de lois antiterroristes ou contre la cybercriminalité (Chak, *op. cit.* ; Maréchal, *op. cit.*). Enfin, des mesures que l'on affirme destinées à contrer l'ingérence étrangère servent à restreindre la liberté d'expression et de réunion d'opposants politiques, contribuant ainsi à délégitimer les mouvements de la société civile (Kogan, *op. cit.*).
- 34 Les écrits soulignent le rôle actif des médias d'État dans la propagande, les restrictions imposées à la liberté de la presse par des cadres réglementaires nationaux et la mise en place d'infrastructures de communication renforçant le contrôle étatique des contenus (Maréchal, *op. cit.* ; Kasm, *op. cit.*). L'intelligence artificielle (IA) est de plus en plus exploitée à des fins de manipulation et de déstabilisation des adversaires politiques, dont les défenseur·es des droits humains, en complément des actions posées par des « armées électroniques » composées de trolls, de censeurs et de bots mobilisés pour harceler, désinformer et saturer l'espace médiatique d'informations trompeuses (Gregory, 2019 ; voir aussi Tenove, *op. cit.*). Pearce et Kendzior (2012), à l'instar de

Gregory (2019), mobilisent le terme « diabolisation » pour désigner une stratégie des régimes autoritaires visant à délégitimer les discours dissidents sur les réseaux sociaux tout en orchestrant des campagnes d'intimidation et de salissage public. Ces pratiques, qui démoralisent et isolent les défenseur·es des droits humains, mettent en cause leur sécurité physique en les rendant sujets à des accusations mensongères exposées publiquement (East and Horn of Africa Human Rights Defenders Project, *op. cit.* ; Michaelsen, *op. cit.*). Enfin, la surveillance ciblée des défenseur·es des droits humains s'accompagne d'intrusions dans leurs communications, de piratages et d'attaques numériques. L'expression *Weaponization of online spaces* est à cet égard fréquemment employée dans la littérature pour décrire cette dynamique (Gregory, 2019, p. 11-12 ; voir aussi Deutch et Al Khatib, *op. cit.* ; Kasm, *op. cit.* ; Tenove, *op. cit.* ; Freeman, *op. cit.*).

- 35 La littérature recense plusieurs pratiques de surveillance étatique, y compris le suivi des activités en ligne des dissidents (Bennett, 2015 ; Dragu et Lupu, 2021 ; Kasm, *op. cit.* ; Nyst, *op. cit.* ; Siatitsa, *op. cit.*), l'interception de données au moyen de logiciels espions (Anstis, *op. cit.* ; Chubb et Roberts Lyer, 2024 ; Eguridu *et al.*, 2024), la traque par GPS et l'extraction des métadonnées (Siatitsa, *op. cit.* ; Dragu et Lupu, *op. cit.*) ainsi que l'usage de la reconnaissance faciale pour identifier les opposants (Gregory, 2019 ; Siatitsa, *op. cit.*). L'utilisation d'outils numériques à des fins militantes favorise la surveillance des individus et des groupes mobilisés qui les emploient : les données générées par ces outils et leurs usages permettent aux autorités de mieux cibler leurs actions et d'identifier leurs adversaires (Dutt et Rasul, *op. cit.*). Cette surveillance, souvent imperceptible (Nyst, *op. cit.*), repose fréquemment sur l'exploitation de failles de sécurité présentes dans les infrastructures et réseaux numériques (Schaake, 2012). Les techniques de piratage et d'infiltration de réseaux activistes mises en évidence dans la littérature témoignent aussi de l'usage de logiciels malveillants et espions par les autorités (Nyst, *op. cit.* ; Siatitsa, *op. cit.*) pour récupérer des informations critiques sur les défenseur·es des droits humains et le déploiement de cyberattaques ciblées — par exemple par déni de service (Marelli, 2020) — pour nuire à leurs activités.

Technologies numériques et défense des droits humains : facteurs conditionnant le potentiel d'usage

- 36 L'intégration de pratiques numériques dans le cadre des activités de défense de droits humains est largement tributaire des facteurs qui conditionnent le potentiel d'usage associé aux technologies numériques. Nos analyses permettent d'identifier les grandes catégories recensées par la littérature à cet égard. Par potentiel d'usage des technologies numériques, nous entendons l'ensemble des éléments affectant les capacités d'individus ou de groupes de la société civile à mobiliser les outils jugés appropriés pour mener leurs activités de défense des droits humains et parvenir à communiquer, à s'organiser et à s'informer dans les espaces numériques. Nos analyses ont fait émerger cinq grandes catégories de facteurs qui, selon les cas, soutiennent ou limitent le potentiel d'usage des technologies numériques à des fins de défense des droits humains. Ces catégories énumèrent des facteurs :
1. d'ordre technique et matériel ;
 2. liés aux capacités des organisations à se saisir des technologies numériques ;
 3. découlant de l'environnement politique et des cadres de réglementation ;
 4. associés à la sécurité numérique ;

5. et, enfin, propres aux caractéristiques de l'écosystème médiatique dans lequel s'inscrivent les pratiques de défense des droits humains.
- 37 L'inventaire des facteurs limitant le potentiel d'usage témoigne d'une nette prédominance d'éléments jugés limitatifs recensés dans la littérature identifiée pour notre étude. Deux fois plus d'articles présentent des facteurs limitatifs (n = 45) que contributifs (n = 23) à l'intégration des technologies numériques dans le cadre de pratiques de défense des droits humains.
- 38 La présence d'infrastructures de télécommunications fiables (Aronson et *al.*, 2018, Sarkin, 2021) et résilientes face aux coupures et au blocage ainsi que l'accès à des technologies numériques de pointe — dans le but de valider l'information audiovisuelle (Freeman, *op. cit.*) —, à des codes propriétaires et sécurisés (Marelli, *op. cit.*), à des réseaux de codeurs et de développeurs travaillant de manière collaborative (Sorell, *op. cit.*) et à des plateformes alternatives de communication (Brough et Li, *op. cit.*) appuient les activités de défense des droits humains. Par opposition, la littérature souligne le fardeau porté par les organisations de défense de droits humains en matière de mise à jour de leur sécurité numérique (Kazansky, *op. cit.* ; Naimer et *al.*, 2017), d'inégalités numériques persistantes, notamment entre hommes et femmes (Shirazi, 2012 ; Dutt et Rasul, *op. cit.*) et entre zones rurales et urbaines (Chak, *op. cit.* ; Khamdamova, *op. cit.*), de ressources financières restreintes (Sarkin, *op. cit.* ; Thrall et *al.*, 2014), d'accès limité à l'imagerie satellite (Herscher, *op. cit.*), d'un manque d'équipements et de formation (Kingston et Stam, *op. cit.*), de coûts élevés (Markland, *op. cit.*), de problèmes de connectivité (Deutch et Al Khatib, *op. cit.*) et d'une surabondance de données pouvant être analysées (Aronson et *al.*, *op. cit.* ; Ristovska, 2016).
- 39 De même, la présence d'informateurs et d'enquêteurs formés aux technologies numériques (Gregory, 2019 ; Tuckwood, 2014), l'accès à des formations consacrées spécifiquement à la sécurité des femmes activistes (Shirazi, *op. cit.*) et une inclusion des organisations dans des réseaux de collaboration visant à développer la littératie et la sécurité numériques (Chak, *op. cit.* ; Deutch et Al Khatib, *op. cit.*) participent au développement des capacités numériques des organisations de défense des droits humains. Le développement de ces capacités est toutefois limité par plusieurs éléments : un manque de littératie numérique et de littératie liée aux données (Aronson et *al.*, *op. cit.* ; Markland, *op. cit.*), la persistance de défis pour participer à des formations en sécurité numérique (tels que le manque de temps, la dépendance à l'égard de formateurs-experts et le peu de ressources à disposition) (Hankey et Ó Clunaigh, *op. cit.* ; Kogan, *op. cit.* ; Schaake, *op. cit.* ; Kazansky, *op. cit.*), un manque d'expertise en analyse et en visualisation des données (Rall et *al.*, 2016) et un conservatisme professionnel à l'égard de l'intégration des technologies numériques au sein des organisations (Markland, *op. cit.*).
- 40 Les politiques publiques et les cadres régulateurs nationaux des communications numériques constituent un important facteur contributif ou limitatif aux usages des technologies numériques en contexte de défense des droits humains. La littérature note que ces usages sont appuyés par des plaidoyers en faveur de l'accès à Internet comme droit fondamental (Joyce, *op. cit.*) et par un renforcement du rôle joué par des organisations internationales dans la protection des libertés numériques (Schaake, *op. cit.*). Des auteurs, à l'instar de Zanzoun (2017), constatent une hésitation (ou un manque de ressources) de certains régimes autoritaires à couper les accès aux réseaux

numériques⁹ et rendent compte d'activités des groupes de la société civile visant à influencer les normes de développement technique et la régulation des communications (Brough et Li, *op. cit.* ; Gregory, 2019).

- 41 On observe une nette prédominance des textes décrivant l'environnement politique et les cadres réglementaires comme des obstacles à l'intégration des technologies numériques dans la défense des droits humains. Les écrits recensés sont près de quatre fois plus nombreux à souligner leurs effets limitatifs (n = 23) plutôt que contributifs (n = 6). Les principales barrières identifiées incluent :

- les risques judiciaires liés à l'usage des technologies numériques à des fins militantes (Maréchal, *op. cit.* ; Onuch, *op. cit.*) ;
- les politiques de censure et de filtrage mises en place par les autorités publiques (Kasm, *op. cit.* ; Dutt et Rasul, *op. cit.*) ;
- les règles arbitraires imposées par les plateformes dont dépendent les groupes militants (Gregory, 2019 ; Kazansky, *op. cit.*) ;
- des mesures insuffisantes ou inadéquates de protection de l'anonymat des usagers des grandes entreprises technologiques (Kingston et Stam, *op. cit.* ; Tenove, *op. cit.*) ;
- le ciblage des proches des défenseur·es des droits humains par les autorités publiques (Mehandru et Koenig, 2019 ; Michaelsen, *op. cit.*) ;
- la manipulation ou la falsification de données essentielles sur les droits humains par les pouvoirs publics (Markland, *op. cit.* ; Sarking, *op. cit.*) ;
- l'absence d'un cadre international cohérent pour lutter contre la surveillance de masse (Dutt et Rasul, *op. cit.*) ;
- la difficulté d'intégrer les normes des droits humains dans l'architecture et les protocoles d'Internet (Cath, 2021 ; Mueller et Badiei, *op. cit.*) ;
- l'incompatibilité de certaines législations nationales en matière de surveillance avec les standards internationaux des droits humains (Nyst, *op. cit.* ; Siatitsa, *op. cit.*) ;
- la lente adoption de standards permettant l'intégration de preuves numériques devant les tribunaux internationaux (Deutch et Al Khatib, *op. cit.* ; Van der Velden, *op. cit.*) ;
- les restrictions légales persistantes entravant le développement d'algorithmes de chiffrement des données (Hankey et Ó Clunaigh, *op. cit.*).

- 42 La question de la sécurité numérique est déterminante pour l'intégration des technologies numériques dans les pratiques de défense des droits humains. La littérature présente toutefois quasi exclusivement cette question sous l'angle des facteurs qui en limitent le potentiel d'usage. La question des menaces numériques y est prépondérante. Ces menaces regroupent le piratage des dispositifs pour récolter des informations privées critiques (Dutt et Rasul, *op. cit.* ; Marelli, *op. cit.*), la cyberintimidation (Michaelsen, *op. cit.* ; Tenove, *op. cit.*), le risque d'identification par reconnaissance faciale (Gregory, 2019 et 2010 ; Dragu et Lupu, *op. cit.*), la diffusion de contenus trompeurs ou falsifiés tels que les hypertrucages (Freeman, *op. cit.* ; Mehandru et Koenig, *op. cit.*), les cyberattaques par déni de service (Hankey et Ó Clunaigh, *op. cit.* ; Sorell, *op. cit.*), le pistage des utilisateurs à l'aide des données de localisation (Kazansky, *op. cit.* ; Van der Velden, *op. cit.*), l'infiltration des réseaux numériques des militants (Kazansky, *op. cit.* ; Tenove, *op. cit.*) ainsi que l'usurpation de l'identité d'une organisation ou l'exploitation de son nom ou de sa réputation (Marelli, *op. cit.*).

- 43 Finalement, les configurations de l'écosystème médiatique conditionnent le potentiel d'usage des technologies numériques à des fins de défense de droits humains en fonction de la nature des relations entretenues par les défenseur·es des droits humains

avec les médias grand public, de leurs accès aux plateformes de réseaux sociaux numériques et des modes de fonctionnement de ces dernières. Les plateformes de réseaux sociaux numériques sont ainsi jugées garantir un accès démocratisé à la diffusion de vidéos, ouvrant également à des occasions de cibler plus efficacement des publics préidentifiés et susceptibles de contribuer à l'établissement de « communautés de confiance » (Gregory, 2010 ; Brough et Li, *op. cit.* ; Kasm, *op. cit.* ; Michaelson, *op. cit.* ; Gregory, 2019). Cependant, soulignons que notre corpus comprend deux fois plus de textes présentant les facteurs limitatifs (n = 12) que contributifs (n = 5) à l'égard de l'écosystème médiatique. La littérature souligne effectivement des difficultés associées aux campagnes de désinformation prenant place dans les environnements numériques (Gregory, 2019), une fragmentation des publics dans les environnements numériques (Chak, *op. cit.* ; Freeman, *op. cit.*), le risque de nourrir des formes jugées peu significatives d'activisme du clic (Joyce, *op. cit.*) ainsi qu'une dépendance persistante à l'égard des médias grand public. Les écrits relèvent également les difficultés à attirer l'attention sur les violations de droits humains dans un contexte marqué par la surabondance de contenus légers et distrayants dans les environnements numériques ainsi que les risques associés à la revictimisation des personnes qui ont vécu des abus ou des violences lorsque des preuves visuelles sont diffusées dans des environnements numériques (Gregory, 2010 ; Ristovska, *op. cit.* ; Van der Velden, *op. cit.*).

Discussion : la mise en données de la défense des droits humains

- 44 L'analyse des écrits permet de valider le constat d'un processus généralisé de numérisation des pratiques constitutives de la défense des droits humains. Tant les architectures que les dispositifs, services et applications numériques jouent un rôle déterminant dans les activités d'enquête et de consignation, de plaidoyer, de mise sous pression des autorités et de soutien aux victimes de violation de droits humains. Le spectre et la diversité des pratiques numériques témoignent de capacités démontrées, chez les défenseur·es des droits humains, à l'innovation, à l'expérimentation et à l'adaptation.
- 45 Dans ce contexte, les processus de mise en données revêtent une importance déterminante dans les pratiques d'enquête et de consignation. Trois éléments fondamentaux structurent la mise en données de ces pratiques. Premièrement, les données relatives aux violations sont activement recherchées, collectées, agrégées et analysées par les défenseur·es des droits humains. Cette collecte peut s'opérer par la production directe de données primaires — avec l'usage de drones, de caméras ou d'imagerie satellite pour « capturer » des preuves — ou par l'exploration et l'exploitation de données secondaires collectées sur les réseaux numériques (Freeman, *op. cit.* ; Gregory, 2019 ; Gutierrez, *op. cit.* ; Herscher, *op. cit.* ; Sarkin, *op. cit.*). Deuxièmement, les défenseur·es des droits humains valorisent ces données en les croisant, en menant des analyses approfondies et en générant de nouvelles données à valeur ajoutée (voir Aronson et al., *op. cit.* ; Tuckwood, *op. cit.*). Troisièmement, ces données sont traitées et mises en forme pour servir des stratégies de plaidoyer et de communication publique. Elles font l'objet d'un travail de vérification et de structuration pour constituer des preuves recevables, compréhensibles et convaincantes des violations constatées (Tuckwood, *op. cit.* ; Kasm, *op. cit.* ; Gregory,

2019). Au-delà de leur fonction probatoire, elles visent aussi à sensibiliser, à émouvoir et à engager divers publics, en mobilisant des récits et des expériences immersives pour susciter adhésion et action en faveur des victimes (Gregory, 2019 ; Gutierrez, *op. cit.*). Les données sont également utilisées pour soutenir et étayer des contre-récits qui s'opposent aux discours assumés par les pouvoirs publics et les invalident, en rectifiant des faits et en illustrant des inexactitudes, des mensonges et des omissions dans ces discours (*ibid.*).

- 46 Toutefois, si nos travaux rendent compte de pratiques d'adaptation et d'innovation technologiques de la part de groupes de défense des droits humains, ils soulignent que celles-ci prennent place dans un contexte où les configurations techniques, le cadre réglementaire et la capacité d'action des États — notamment en matière de surveillance et de répression numériques — s'imposent à eux. Les pratiques numériques de ces groupes sont ainsi influencées par des conditions structurelles sur lesquelles ils n'exercent qu'une influence indirecte et par des capacités internes limitées, en termes de compétences ainsi que de ressources humaines, financières et techniques (Rall et *al.*, *op. cit.*). Les pratiques numériques des défenseur·e·s des droits humains doivent ainsi être contextualisées et étudiées dans un cadre marqué par des rapports de pouvoir asymétriques à l'avantage de leurs adversaires. Comme le soulignent Dragu et Lupu, les gouvernements ont

utilisé la reconnaissance vocale pour scanner les réseaux mobiles, suivi les déplacements des citoyens par GPS, lu des courriels et des messages texte pour surveiller les groupes dissidents et censurer sélectivement certaines informations. Ils ont également recours à des logiciels malveillants et à des logiciels espions pour activer secrètement les webcams intégrées aux ordinateurs portables personnels et les microphones des téléphones portables. L'Internet, en particulier, facilite l'utilisation de nombreux outils au service de ces gouvernements, notamment des dispositifs sophistiqués de surveillance numérique. Nous appelons cet effet de la technologie « l'effet de contrôle préventif » (*op. cit.*, p. 6, traduction libre).

- 47 Les processus de numérisation de la défense des droits humains, et plus spécifiquement le recours quotidien et difficilement évitable à des technologies qui regroupent et mettent en données des informations, représentent des vulnérabilités importantes pour les défenseur·e·s des droits humains. L'atténuation des risques et des dépendances numériques à l'égard d'acteurs externes constitue conséquemment un enjeu déterminant pour l'évolution des pratiques numériques des défenseur·e·s des droits humains.
- 48 Nos travaux mettent en évidence quatre types de mesures adoptées par les défenseur·e·s des droits humains à cette fin. Le premier priorise le développement de la littératie numérique et des compétences en gestion des données chez les individus et les groupes engagés dans la défense des droits humains (Chak, *op. cit.*). Le deuxième repose sur le renforcement des capacités techniques autonomes par le biais de la création et de l'utilisation d'infrastructures, de dispositifs et de services numériques alternatifs¹⁰ et non commerciaux (Chak, *op. cit.* ; Gregory, 2019). Le troisième vise à exercer une pression sur les grandes entreprises du numérique pour qu'elles intègrent des principes de protection des droits humains dans leurs pratiques et veillent à se conformer aux meilleurs standards d'intégration des droits humains dans les processus de conception de leurs outils et services technologiques (Gregory, 2019 ; Mueller et Badiei, *op. cit.*). Enfin, les défenseur·e·s des droits humains interpellent les instances réglementaires et judiciaires pour favoriser un environnement numérique plus sûr et

inclusif pour la militance démocratique (Gregory, 2019) ainsi que des mises à jour des normes d'inclusion des preuves numériques dans le système judiciaire (Deutch et Al Khatib, *op. cit.*).

Conclusion

- 49 Cet article a procédé à une recension des écrits savants sur des pratiques numériques de défense et de répression des droits humains. À ce titre, il reflète nécessairement les préférences et priorités exprimées par la communauté universitaire de langue anglaise concernant ces pratiques, sans pour autant permettre de corriger les angles morts de la recherche. Plus de la moitié des premiers auteurs et autrices des articles constituant le corpus viennent de trois pays anglo-saxons — les États-Unis (n = 16), la Grande-Bretagne (n = 8) et le Canada (n = 4) —, ce qui influence la perspective adoptée¹¹. De même, la focale appliquée à la littérature de langue anglaise permet de recenser les perspectives universitaires les plus diffusées à l'échelle internationale, l'anglais constituant la principale langue de publication. Toutefois, elle entraîne inévitablement l'exclusion de travaux publiés dans d'autres langues qui abordent des enjeux, des thèmes, des pratiques, des acteurs et des contextes sociaux, culturels et politiques pouvant s'avérer parfaitement pertinents. Par ailleurs, la période couverte (2010-2021) a eu un effet déterminant sur les résultats de l'étude en exposant de pratiques historiquement situées, inscrites dans des contextes technologiques, culturels, sociaux et politiques spécifiques.
- 50 Cette décennie a été marquée par une accélération du développement technologique, la diffusion massive de dispositifs, services et plateformes numériques ainsi que par des mobilisations sociales d'envergure reposant sur des réseaux de communication numérique. Elle a également été le théâtre d'une prise de conscience internationale face à l'emprise croissante des États et des GAFAM sur les architectures et infrastructures numériques de communication. Cette période a effectivement vu l'émergence d'un oligopole numérique transnational dominé par quelques acteurs peu réglementés, entretenant des relations étroites avec les États. Elle s'achève avec le déploiement de technologies d'IA perturbatrices, dont les effets directs et significatifs soulèvent d'importants enjeux en matière de défense des droits humains.
- 51 Cette décennie met ainsi en lumière les manières avec lesquelles l'extraction, la circulation, la modélisation et la gouvernance des données transforment des pratiques démocratiques ou répressives bien établies. Elle illustre une mise en service des données au profit de stratégies sociopolitiques, tant dans le cadre de stratégies de contrôle et de surveillance (Dijck, *op. cit.*) que d'initiatives de résistance politique (Milan et Van der Velden, *op. cit.*). Des travaux sont en cours pour couvrir une période plus récente (2022-2024), marquée par la diffusion à grande échelle de technologies d'IA, une consolidation des oligopoles numériques, des alliances renouvelées entre des acteurs privés et des États — surtout à des fins de surveillance et de contrôle — et une répression accrue des défenseur·es des droits humains sur la scène internationale.
- 52 Les premiers résultats signalent une augmentation substantielle du volume d'écrits qui se consacrent aux pratiques numériques de défense et de répression des droits humains ainsi qu'une politisation marquée des données. Les travaux dont il est question dans le présent article témoignent des tendances, des contextes et des évolutions qui orientent un phénomène désormais solidement constitué d'une instrumentalisation croissante

des données, de nature tactique et stratégique, dans le cadre de conflits de violation de droits humains.

BIBLIOGRAPHIE

- AGENCE DE SANTÉ PUBLIQUE DU CANADA (2007), *Les jeunes de la rue au Canada et l'utilisation de substances : constatations découlant de la surveillance accrue des jeunes de la rue au Canada, 1999-2003*, Ottawa, L'Agence, https://publications.gc.ca/collections/collection_2008/phac-aspc/HP5-23-2007F.pdf, page consultée le 9 mars 2016.
- AL-RAWI Ahmed (2021), « Disinformation under a networked authoritarian state: Saudi Trolls' credibility attacks against Jamal Khashoggi », *Open Information Science*, 5(1), p. 140-162, DOI : <https://doi.org/10.1515/opis-2020-0118>.
- AL-RAWI Ahmed, Chris TENOVE et Peter KLEIN (2023), « The networked trolling of critical journalists and news organizations in Iraq », *Journalism and Media*, 4(4), p. 1130-1140, <https://www.mdpi.com/2673-5172/4/4/72>.
- ANSTIS Siena (2024), « Regulating transnational dissident cyber espionage », *International & Comparative Law Quarterly*, 73(1), p. 259-274, <https://www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/regulating-transnational-dissident-cyber-espionage/8662095ACD8DB0BB32392E1BAD7DEFF6>, page consultée le 13 juillet 2025.
- ARONSON Jay D., McKenna COLE, Alex HAUPTMANN, Dan MILLER et Bradley SAMUELS (2018), « Reconstructing human rights violations using large eyewitness video collections: The case of Euromaidan protester deaths », *Journal of Human Rights Practice*, 10(1), p. 159-178, DOI : <https://doi.org/10.1093/jhuman/huy005>.
- BENNETT Karen (2015), « European Union guidelines on human rights defenders: A review of policy and practice towards effective implementation », *The International Journal of Human Rights*, 19(7), p. 908-934, DOI : <https://doi.org/10.1080/13642987.2015.1075303>.
- BENESSAIEH Karim (2009), « Lock-out au Journal de Montréal. Quebecor durcit ses demandes », *La Presse*, 16 décembre, p. A14.
- BHARGAVA Rahul, Erica DE AHL, Amanda LETOUZÉ, Amanda NOONAN, David SANGOKOYA et Natalie SHOUP (2015), « Beyond data literacy: Reinventing community engagement and empowerment in the age of data », *Data-Pop Alliance White Paper Series*, Data-Pop Alliance, Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute, and Internews, septembre.
- BRESSON Maryse (2008), « La participation : un concept constamment réinventé. Analyse sociologique des enjeux de son usage et de ses variations », *Sociologies. Revue de l'Association française de sociologie*, <http://socio-logos.revues.org/2817>, page consultée le 4 décembre 2015.
- BROUGH Melissa et Zhan LI (2013), « Media systems dependency and human rights online video: The "Saffron Revolution" and WITNESS's Hub », *International Journal of Communication*, 7, <https://ijoc.org/index.php/ijoc/article/view/1423>, page consultée le 16 novembre 2025.

CASEMAJOR Nathalie (2023), « Cultures de données et publics de données : conceptualisation critique », dans Frédéric MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Datafication et publics de données. Penser la mise en donnée de la société*, Montréal, Presses de l'Université de Montréal, <https://hal.science/hal-04254885>, page consultée le 16 novembre 2025.

CASCIO M. Ariel, Eunlye LEE, Nicole VAUDRIN et Darcy A. FREEDMAN (2019), « A team-based approach to open coding: Considerations for creating intercoder consensus », *Field Methods*, 31(2), p. 116-130, DOI : <https://doi.org/10.1177/1525822X19838237>.

CATH Corinne (2021), « The technology we choose to create: Human rights advocacy in the Internet Engineering Task Force », *Telecommunications Policy*, 45(6), DOI : <https://doi.org/10.1016/j.telpol.2021.102144>

DE CERTEAU Michel (1984), *The practice of everyday life*, Berkeley, University of California Press.

CHAK Sopheap (2014), « New information and communication technologies' influence on activism in Cambodia human rights in motion: Tools: Essay », *Sur - International Journal on Human Rights*, 20, p. 437-448, <https://heinonline.org/HOL/P?h=hein.journals/surij20&i=437>, page consultée le 16 novembre 2025.

CHIWESHE Manase Kudzai et Primrose HOVE (2023), « Intersectional analysis of women human rights defenders' lived experiences under Covid-19 lockdowns in Zimbabwe », *Gender & Development*, 31(1), p. 89-107, DOI : <https://doi.org/10.1080/13552074.2023.2181539>.

CHUBB Andrew et Kiran ROBERTS LYER (2024), « Transnational human rights violations: Addressing the evolution of globalized repression through national human rights institutions », *Journal of Human Rights Practice*, 16(3), p. 770-793, DOI : <https://doi.org/10.1093/jhuman/huae017>.

COSTANZA-CHOCK Sasha (2020), *Design justice: Community-led practices to build the worlds we need*, Cambridge, The MIT Press, coll. « Information Policy ».

DENOUEL Julie et Fabien GRANJON (dir.) (2011), *Communiquer à l'ère numérique. Regards croisés sur la sociologie des usages*, Paris, Presses des Mines, coll. « Sciences sociales », <https://books.openedition.org/pressesmines/387>, page consultée le 23 novembre 2025.

DEUTCH Jeff et Hadi Habal Al KHATIB (2018), « The Syrian archive: A methodological case study of open-source investigation of state crime using video evidence from social media platforms », *State Crime Journal*, 7(1), DOI : <https://doi.org/10.13169/statecrime.7.1.0046>.

DIJCK Jose van (2014), « Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology », *Surveillance & Society*, 12(2), p. 197-208, DOI : <https://doi.org/10.24908/ss.v12i2.4776>.

DRAGU Tiberiu et Yonatan LUPU (2021), « Digital authoritarianism and the future of human rights », *International Organization*, 75(4), p. 991-1017, DOI : <https://doi.org/10.1017/S0020818320000624>.

DUTT Mallika et Nadia RASUL (2014), « Raising digital consciousness: An analysis of the opportunities and risks facing human rights activists in a Digital Age », *Sur*, 11(20), p. 426-435.

EAST AND HORN OF AFRICA HUMAN RIGHTS DEFENDERS PROJECT (2013), « Networks for the protection of human rights defenders: Notes from the field », *Journal of Human Rights Practice*, 5(3), p. 522-534, DOI : <https://doi.org/10.1093/jhuman/hut027>.

EGURIDU Henry, Tolani TOLANI, Olusola OLUSOLA et Wasiu SAKPERE (2024), « Impact of information technology on civil rights and criminal liabilities », *African Journal of Law, Ethics and*

Education, 7(1), <https://ajleejournal.com/index.php/ajlee/article/view/168>, page consultée le 18 novembre 2025.

FELDSTEIN Steven (2021), *The rise of digital repression: How technology is reshaping power, politics, and resistance*, Oxford, Oxford University Press, DOI : <https://doi.org/10.1093/oso/9780190057497.001.0001>.

FLETCHER Laurel E. (2022), « Power and the international human rights imaginary: A critique of practice », *Journal of Human Rights Practice*, 14(3), p. 749-768, DOI : <https://doi.org/10.1093/jhuman/huac051>.

FREEMAN Lindsay (2021), « Weapons of war, tools of justice: Using artificial intelligence to investigate international crimes », *Journal of International Criminal Justice*, 19(1), p. 35-53, DOI : <https://doi.org/10.1093/jicj/mqab013>.

GARCIA Amelia A., Michael J. BRITTON, Dev D. DOSHI, Munmun DE CHOUDHURY et Carl A. LE DANTEC (2020), « Data migrations: Exploring the use of social media data as evidence for human rights advocacy », *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), p. 1-25, <https://ledantec.net/wp-content/uploads/2020/09/CSCW-Data-Migrations.pdf>, page consultée le 18 novembre 2025.

GIESEN Lindsay et Allison ROESER (2020), « Structuring a team-based approach to coding qualitative data », *International Journal of Qualitative Methods*, 19, DOI : <https://doi.org/10.1177/1609406920968700>.

GRANJON Fabien (2017), « Résistances en ligne : mobilisation, émotion, identité », *Variations. Revue internationale de théorie critique*, 20, DOI : <https://doi.org/10.4000/variations.819>.

GRANJON Fabien et Julie DENOÛËL (2011), « Présentation. Penser les usages sociaux des technologies numériques d'information et de communication », dans *Communiquer à l'ère numérique*, Paris, Presses des Mines, coll. « Sciences sociales », p. 7-43.

GREGORY Sam (2010), « Cameras everywhere: Ubiquitous video documentation of human rights, new forms of video advocacy, and considerations of safety, security, dignity and consent », *Journal of Human Rights Practice*, 2(2), p. 191-207, DOI : <https://doi.org/10.1093/jhuman/huq002>.

GREGORY Sam (2019), « Cameras everywhere revisited: How digital technologies and social media aid and inhibit human rights documentation and advocacy », *Journal of Human Rights Practice*, 11(2), p. 373-392, DOI : <https://doi.org/10.1093/jhuman/huz022>.

GUTIERREZ Miren (2021), « Data activism and meta-documentary in six films by Forensic Architecture », *Studies in Documentary Film*, 17(1), p. 32-52, DOI : <https://doi.org/10.1080/17503280.2021.1908932>.

HANKEY Stephanie et Daniel Ó CLUNAIGH (2013), « Rethinking risk and security of human rights defenders in the Digital Age », *Journal of Human Rights Practice*, 5(3), p. 535-547, DOI : <https://doi.org/10.1093/jhuman/hut023>.

HERSCHER Andrew (2014), « Surveillant witnessing: Satellite imagery and the visual politics of human rights », *Public Culture*, 26(3), p. 469-500, DOI : <https://doi.org/10.1215/08992363-2683639>.

HOLCOMB Jesse, Kim GROSS et Amy MITCHELL (2011), *How mainstream media outlets use Twitter: Content analysis shows an evolving relationship*, Washington (DC), School of Media and Public Affairs, <http://www.journalism.org/sites/journalism.org/files/How%20Mainstream%20Media%20Outlets%20UseTwitter.pdf>, page consultée le 24 novembre 2011.

- HOWARD Philip N., Sheetal D. AGARWAL et Muzammil M. HUSSAIN (2011), « When do States disconnect their digital networks? Regime responses to the political uses of social media », *APSA 2011 Annual Meeting Paper*, DOI : <http://dx.doi.org/10.2139/ssrn.1907191>.
- JOYCE Daniel (2015), « Internet freedom and human rights », *European Journal of International Law*, 26(2), p. 493-514, DOI : <https://doi.org/10.1093/ejil/chv021>.
- KANE Oumar (2013), « Les usages des TIC entre analyse sociotechnique et théories de l'appropriation : état de la littérature », *Les Enjeux de la communication*, janvier, p. 23-42, <https://www.academia.edu/7277534>, page consultée le 18 novembre 2025.
- KASM Null (2018), « Redefining publics: Mosireen, state crime and the rise of a digital public sphere », *State Crime Journal*, avril, DOI : <https://doi.org/10.13169/statecrime.7.1.0100>.
- KAUN Anne et Emiliano TRERÉ (2020), « Repression, resistance and lifestyle : charting (dis)connection and activism in times of accelerated capitalism », *Social Movement Studies*, 19(5-6), p. 697-715, DOI : <https://doi.org/10.1080/14742837.2018.1555752>.
- KAZANSKY Becky (2015), « FCJ-195 Privacy, responsibility, and human rights activism », *The Fibreculture Journal*, 26, p. 190-208, DOI : <https://doi.org/10.15307/fcj.26.195.2015>.
- KHAMDAMOVA Firuza (2020), « The Impact of digital technologies on human rights », *TSUL Legal Report International Electronic Scientific Journal*, 1(1), <https://legalreport.tsul.uz/index.php/journal/article/view/21>, page consultée le 18 novembre 2025.
- KINGSTON Lindsey N. et Kathryn R. STAM (2013), « Online advocacy: Analysis of human rights ngo websites », *Journal of Human Rights Practice*, 5(1), p. 75-95, DOI : <https://doi.org/10.1093/jhuman/hus036>.
- KOGAN Vanessa (2013), « Protecting human rights defenders in the North Caucasus: Reflections on developments from 2009 to the present », *Journal of Human Rights Practice*, 5(3), p. 500-511, DOI : <https://doi.org/10.1093/jhuman/hut015>.
- LAND Molly B. (2009), « Networked activism », *Harvard Human Rights Journal*, 22, p. 205-243, https://www.mobileactive.org/files/file_uploads/Networked%20Activism.pdf, page consultée le 01 décembre 2025.
- LAND Molly K. et Jay D. ARONSON (2020), « Human rights and technology: New challenges for justice and accountability », *Annual Review of Law and Social Science*, 16, p. 223-240, DOI : <https://doi.org/10.1146/annurev-lawsocsci-060220-081955>.
- LANDRY Normand, France AUBIN et Michel SÉNÉCAL (2015), « Médias et mouvements sociaux : pratiques de mobilisations collectives », *Canadian Journal of Media Studies - Revue canadienne d'études médiatiques*, <https://hal.science/hal-01405451/>, page consultée le 18 novembre 2025.
- LIEVROUW Leah A. (2023), *Alternative and Activist New Media*, 2^e éd., Cambridge, Polity, coll. « Digital Media and Society Series ».
- LOCKE Karen, Martha FELDMAN et Karen GOLDEN-BIDDLE (2022), « Coding practices and iterativity: Beyond templates for analyzing qualitative data », *Organizational Research Methods*, 25(2), p. 262-284, DOI : <https://doi.org/10.1177/1094428120948600>.
- LYON David (2015), *Surveillance after Snowden*, Cambridge, Malden (MA), Polity Press.
- MAKHORTYKH Mykola (2023), « Unreliable narrators or untimely archivists? Challenges of using digital platforms for documenting and remembering Russia's War in Ukraine », *Georgetown Journal of International Affairs*, 24(2), p. 165-173, DOI : <https://doi.org/10.1353/gia.2023.a913642>.

- MARÉCHAL Nathalie (2015), « Ranking digital rights: Keeping the Internet safe for advocacy », *The Fibreculture Journal*, 26, p. 302-308, DOI : <https://doi.org/10.15307/fcj.mesh.009.2015>.
- MARELLI Massimo (2020), « Hacking humanitarians: Defining the cyber perimeter and developing a cyber security strategy for international humanitarian organizations in digital transformation », *International Review of the Red Cross*, 102(913), p. 367-387, DOI : <https://doi.org/10.1017/S1816383121000151>.
- MARKLAND Alistair (2021), « Epistemic transformation at the margins: Resistance to digitalisation and datafication within global human rights advocacy », *Global Society*, 36(1), p. 113-133, DOI : <https://doi.org/10.1080/13600826.2021.1879028>.
- McNUTT John G. (dir.) (2018), *Technology, activism, and social justice in a Digital Age*, Oxford, Oxford University Press.
- McPHERSON Ella (2015), « Surmounting the verification barrier between the field of professional human rights: Fact-finding and the non-field of digital civilian witnessing », dans Rebecca Ann LIND (dir.), *Produsing Theory 2.0: The Intersection of Audiences and Production in a Digital World*, New York, Peter Lang Publishing, p. 193-210.
- MEHANDRU Nikita et Alexa KOENIG (2019), « ICTs, social media, & the future of human rights », *Duke Law & Technology Review*, 17(1), p. 129-145, <https://scholarship.law.duke.edu/dltr/vol17/iss1/5>, page consultée le 18 novembre 2025.
- MEIKLE Graham (2018), *The Routledge companion to media and activism*, Londres, New York, Routledge, coll. « Routledge Media and Cultural Studies Companions ».
- MICHAELSEN Marcus (2018), « Exit and voice in a Digital Age: Iran's exiled activists and the authoritarian state », *Globalizations*, 15(2), p. 248-264, DOI : <https://doi.org/10.1080/14747731.2016.1263078>.
- MILAN Stefania (2013), *Social movements and their technologies*, Londres, Palgrave Macmillan UK, DOI : <https://doi.org/10.1057/9781137313546>.
- MILAN Stefania et Lonneke VAN DER VELDEN (2016), « The alternative epistemologies of data activism », *Digital Culture & Society*, 2(2), p. 57-74, DOI : <https://doi.org/10.14361/dcs-2016-0205>.
- MOROZOV Evgeny (2011), *The Net delusion: The dark side of Internet freedom*, New York, PublicAffairs.
- MUELLER Milton L. et Farzaneh BADIEI (2019), « Requiem for a dream: On advancing human rights via Internet architecture », *Policy & Internet*, 11(1), p. 61-83, DOI : <https://doi.org/10.1002/poi3.190>.
- NAH Alice M., Karen BENNETT, Danna INGLETON et James SAVAGE (2013), « A research agenda for the protection of human rights defenders », *Journal of Human Rights Practice*, 5(3), p. 401-420, DOI : <https://doi.org/10.1093/jhuman/hut026>.
- NAIMER Karen, Widney BROWN, Ranit MISHORI et PHYSICIANS FOR HUMAN RIGHTS (2017), « MediCapt in the Democratic Republic of the Congo: The design, development, and deployment of mobile technology to document forensic evidence of sexual violence », *Genocide Studies and Prevention*, 11(1), p. 25-35, DOI : <https://doi.org/10.5038/1911-9933.11.1.1455>.
- NYST Carly (2018), « Secrets and lies: The proliferation of state surveillance capabilities and the legislative secrecy which fortifies them – an Activist's account », *State Crime Journal*, 7(1), p. 8-23, DOI : <https://doi.org/10.13169/statecrime.7.1.0008>.

OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE (2014), « Mise en données », *Vitrine linguistique*, <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26523020/mise-en-donnees>, page consultée le 10 avril 2025.

ONU (1999), *Déclaration sur le droit et la responsabilité des individus, groupes et organes de la société de promouvoir et protéger les droits de l'homme et les libertés fondamentales universellement reconnus* (Résolution A/RES/53/144, adoptée le 8 mars 1999), New York, Organisation des Nations Unies, <https://www.ohchr.org/fr/instruments-mechanisms/instruments/declaration-human-rights-defenders>, page consultée le 10 avril 2025.

ONUOLGA Olga (2015), « “Facebook Helped Me Do It”: Understanding the EuroMaidan protester “Tool-Kit” », *Studies in Ethnicity and Nationalism*, 15(1), p. 170-184, DOI : <https://doi.org/10.1111/sena.12129>.

PAUL Jeffrey W., Renato BEZERRA RODRIGUES et Jillian SENIUK CICEK (2021), « Pair-coding as a method to support intercoder agreement in qualitative research », dans *2021 IEEE Frontiers in Education Conference (FIE)*, p. 1-5, DOI : <https://doi.org/10.1109/FIE49875.2021.9637248>.

PEARCE Katy E. et Sarah KENDZIOR (2012), « Networked authoritarianism and social media in Azerbaijan », *Journal of Communication*, 62(2), p. 283-298, DOI : <https://doi.org/10.1111/j.1460-2466.2012.01633.x>.

RALL Katharina, Margaret L. SATTERTHWAITE, Anshul Vikram PANDEY, John EMERSON, Jeremy BOY, Oded NOV et Enrico BERTINI (2016), « Data visualization for human rights advocacy », *Journal of Human Rights Practice*, 8(2), p. 171-197, DOI : <https://doi.org/10.1093/jhuman/huw011>.

RISTOVSKA Sandra (2016), « The rise of eyewitness video and its implications for human rights: Conceptual and methodological approaches », *Journal of Human Rights*, 15(3), p. 347-360, DOI : <https://doi.org/10.1080/14754835.2015.1132157>.

RUDIE JD., Zach KATZ, Sam KUH BANDER et Suman BHUNIA (2021), « Technical analysis of the NSO Group's Pegasus Spyware », dans *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, p. 747-752, DOI : <https://doi.org/10.1109/CSCI54926.2021.00188>.

S. A. (2014), « Expérience de Milgram », *Wikipédia, l'encyclopédie libre*, http://fr.wikipedia.org/wiki/Experience_de_Milgram, page consultée le 9 décembre 2014.

SABER Dima et Abdulrahman RAHMAN AL-JALOUD (2024), « The Syrian Archive Digital Memory Project: Archiving as testimony, as evidence, as creative practice », *Visual Anthropology*, 37(1), p. 19-34, DOI : <https://doi.org/10.1080/08949468.2023.2285215>.

SARKIN Jeremy J. (2021), « Promoting the rights of victims in under-resourced places by using science and technology that can be used by ordinary people, to deal with human rights violations: Bolstering the right to the truth », *Liverpool Law Review*, 42(3), p. 339-360, DOI : <https://doi.org/10.1007/s10991-021-09276-1>.

SCHAAKE Marietje (2012), « Digital freedom and security in a globally connected World: Lessons from the MENA region view from the MENA region and partnering organisations », *Security and Human Rights*, 23(3), p. 249-254, <https://heinonline.org/HOL/P?h=hein.journals/helsnk23&i=255>, page consultée le 18 novembre 2025.

SHIRAZI Farid (2012), « Information and communication technology and women empowerment in Iran », *Telematics and Informatics*, 29(1), p. 45-55, DOI : <https://doi.org/10.1016/j.tele.2011.02.001>.

- SIATITSA Ilia (2020), « Freedom of assembly under attack: General and indiscriminate surveillance and interference with Internet communications », *International Review of the Red Cross*, 102(913), p. 181-198, DOI : <https://doi.org/10.1017/S1816383121000047>.
- SIMARD-HOUE Mélodie (2015), *Le reporter, médiateur, écrivain et héros. Un répertoire culture (1870-1939)*, thèse de doctorat en études littéraires sous la direction de Guillaume PINSON et Marie-Ève THÉRENTY, Québec/Montpellier, Université Laval/Université Paul Valéry, <http://www.theses.ulaval.ca/2015/31764/>, page consultée le 8 mars 2016.
- SORELL Tom (2015), « Human rights and hacktivism: The cases of Wikileaks and Anonymous », *Journal of Human Rights Practice*, 7(3), p. 391-410, DOI : <https://doi.org/10.1093/jhuman/huv012>.
- STEFANOVSKA Vesna (2024), « New technologies and enforced disappearances: Opportunities and challenges for the protection of human rights », *Journal of Criminology and Criminal Law (JCCL)*, 62(1), p. 107-128, DOI : <https://doi.org/10.47152/rkkp.62.1.6>.
- TENOVE Christopher (2019), « Networking justice: Digitally-enabled engagement in transitional justice by the Syrian diaspora », *Ethnic and Racial Studies*, 42(11), p. 1950-1969, DOI : <https://doi.org/10.1080/01419870.2019.1569702>.
- THRALL A. Trevor, Dominik STECULA et Diana SWEET (2014), « May we have your attention please? Human-rights ngos and the problem of global communication », *The International Journal of Press/Politics*, 19(2), p. 135-159, DOI : <https://doi.org/10.1177/1940161213519132>.
- TILLY Charles (1995), *Popular contention in Great Britain, 1758-1834*, Cambridge, Harvard University Press.
- TILLY Charles (2008), *Contentious Performances*, Cambridge, Cambridge University Press.
- TILLY Charles et Sidney TARROW (2015), *Contentious politics*, 2^e éd., Oxford, Oxford University Press.
- TUCKWOOD Christopher (2014), « The state of the field: Technology for atrocity response », *Genocide Studies and Prevention: An International Journal*, 8(3), DOI : <https://doi.org/10.5038/1911-9933.8.3.7>.
- VANACKER Lukas (2014), « Populist MEPs lead Twitter ranking », *EU Observer*, 15 mai, <http://euobserver.com/eu-elections/124095>, page consultée le 4 décembre 2015.
- VELDEN Lonneke van der (2015), « Forensic devices for activism: Metadata tracking and public proof », *Big Data & Society*, 2(2), DOI : <https://doi.org/10.1177/2053951715612823>.
- VET Freek van der et Laura LYYTIKÄINEN (2015), « Violence and human rights in Russia: How human rights defenders develop their tactics in the face of danger, 200-2013 », *The International Journal of Human Rights*, 19(7), p. 979-998, DOI : <https://doi.org/10.1080/13642987.2015.1075306>.
- ZANZOUN Mustapha (2017), « The use of social media in promoting human rights among speakers of English in Morocco: The case of Facebook », *International Journal of Education and Research*, 5(9), p. 169-184.
- ZICCARDI Giovanni (2012), *Resistance, liberation technology and human rights in the Digital Age*, Dordrecht, Springer Science & Business Media.
- ZUBOFF Shoshana (2019), *The age of surveillance capitalism: The fight for a human future at the new frontier of power*, New York, PublicAffairs.

NOTES

1. Il convient de préciser ici que la nature des travaux réalisés — reposant sur une analyse systématique de la littérature savante — expose des pratiques recensées par des acteurs universitaires, mais ne peut évidemment pas prétendre appréhender de manière exhaustive l'ensemble des pratiques déployées sur le terrain ni refléter fidèlement la considération que leur accorde la communauté internationale des défenseur·e·s des droits humains. Néanmoins, ces travaux offrent une première et nécessaire entrée pour le recensement et la catégorisation des pratiques numériques étudiées, tout en permettant d'organiser et de structurer les discours universitaires à leur sujet.
2. En droit international, l'expression « droits humains » est principalement traduite en français par « droits de l'homme », bien qu'une tendance s'accélère pour un usage plus étendu du terme « droits humains » qui élimine le biais historique de genre que l'on retrouve dans « droits de l'homme ». Dans le cadre de nos travaux, nous reprenons l'expression neutre en matière de genre partout, sauf en cas de citation de documents officiels.
3. Il existe un large consensus selon lequel les actions collectives menées dans le cadre de la défense des droits humains doivent demeurer non violentes, ne pas porter atteinte de manière significative à autrui et ne constituer en aucun cas une menace à l'intégrité physique ou psychologique des personnes. Toutefois, ce consensus n'exclut pas l'existence d'initiatives qui s'en écartent, comme en témoignent les travaux de Sorell (2015).
4. Pensons, notamment, à l'usage tactique d'attaques du type « déni-de-service », à l'accès aux données confidentielles et à la divulgation de celles-ci, au détournement de sites Web, à l'usage des téléphones intelligents à des fins de coordination pendant les manifestations et à l'usage de ces mêmes appareils pour enregistrer et diffuser des abus et des malversations.
5. HeinOnline — Law Journal Library, Brill Online Books, Human Rights and Humanitarian Law, Oxford scholarship Online — Law, Scopus (Elsevier), CANLII, ProQuest, Google Scholar.
6. Les mots-clés suivants ont été employés pour affiner la collecte : « Human Rights Defender » OU « Human Rights Defenders » OU « Human Rights Activist » OR « Human Rights Activists » ET « Digital Technology » OU « Media Technology » OU « Communication Technology » OU « Digital Technologies » OU « Media Technologies » OU « Communication Technologies ».
7. *International Journal on Human Rights*, *Human Rights Quarterly*, *The International Human Rights Law Review*, *Asia-Pacific Journal on Human Rights and the Law*, *Journal of Human Rights Practice*, *International Review of the Red Cross*.
8. Si nos travaux se concentrent sur les actions des États, les défenseur·e·s des droits humains sont également la cible de pratiques numériques émanant de groupes citoyens opposés, d'intérêts particuliers ou encore d'entreprises privées. Par ailleurs, les pratiques numériques répressives mises en œuvre par les États reposent fréquemment sur la collaboration (voire la complicité) d'acteurs privés.
9. Et ce, en raison d'une volonté de continuer à surveiller les usages à des fins de répression préventive ou encore de ne pas susciter davantage de contestation par le biais de restrictions abusives.
10. Bien que dans le cas des réseaux sociaux numériques alternatifs, le niveau d'adoption demeure insuffisant dans la population (Kazansky, *op. cit.*).
11. La littérature anglo-saxonne, dominante sur la scène internationale, est couramment associée à une approche libérale des droits humains pouvant marginaliser des perspectives plus sociales ou communautaires. Elle met généralement l'accent sur le droit, les processus judiciaires et les organisations formellement constituées — parfois au détriment de regroupements plus informels — et privilégie des thèmes et enjeux susceptibles de reléguer au second plan des perspectives locales ou émergentes.

RÉSUMÉS

Une abondante littérature témoigne de l'intégration progressive des technologies numériques dans les activités des défenseur·e·s des droits humains (DDH). Cet article présente les résultats d'une revue systématique de la littérature sur plus de dix ans (2010-2021) portant sur les pratiques numériques de défense et de répression des droits humains. L'analyse de la littérature met en évidence des répertoires de pratiques numériques à la fois dans la défense des droits humains et dans la répression des personnes et des organismes qui s'y consacrent. Elle permet par ailleurs de dresser un inventaire des facteurs qui exercent une influence sur la capacité des défenseur·e·s des droits humains à mobiliser les technologies numériques dans le cadre de leurs activités. Les constats indiquent que les processus de mise en données sont déterminants dans les pratiques des défenseur·e·s des droits humains et des États impliqués dans leur répression. Ces processus jouent également un rôle crucial dans la détermination des possibilités et des risques des technologiques en contexte de militance démocratique. L'analyse de la littérature met en évidence des représentations et des actions politiques entreprises par les défenseur·e·s des droits humains pour exercer une pression sur les processus réglementaires et de gouvernance des données numériques dans le but de réduire les risques et d'accroître leur capacité d'action.

A wealth of literature attests to the gradual integration of digital technologies in the activities of human rights advocates. This article presents the findings of a systematic review of the literature over more than ten years (2010 to 2021) on digital practices for the protection and repression of human rights. The review highlights repositories of digital practices in both the protection of human rights and the repression of the people and organizations who serve as human rights advocates. It also provides an inventory of factors that affect the extent to which the protectors of human rights can use digital technologies in their activities. The findings indicate that datafication processes are a critical component of the practices of human rights advocates and of the states involved in their repression. These processes also play a crucial role in identifying technology-related risks and opportunities in the context of democratic activism. The literature review highlights political representations and actions taken by human rights advocates to exert pressure on regulatory and digital data governance processes in order to reduce risks and increase their capacity to act.

Existen abundantes escritos que dan testimonio de la integración progresiva de las tecnologías digitales en las actividades de los Defensores de los Derechos Humanos (DDH). Este artículo presenta los resultados del análisis sistemático de los escritos en más de diez años (2010-2021) concernientes a las prácticas digitales de defensa y represión de los derechos humanos. El análisis de los escritos pone en evidencia repertorios de prácticas digitales a la vez sobre la defensa de los derechos humanos, así como, lo referente a la represión en las personas y organismos que se dedican a ello. Por otra parte, esto permite realizar un inventario de los factores que ejercen una influencia en la capacidad que tienen los defensores de los derechos humanos de recurrir a las tecnologías digitales en el marco de sus actividades. Las observaciones indican que los procesos de entrada de datos son determinantes en las prácticas de los defensores de los derechos humanos y los Estados implicados en la represión. También desempeñan un papel importante en la determinación de los riesgos y oportunidades ofrecidas por las tecnologías de militancia democrática. El análisis de los escritos resalta las representaciones y acciones políticas emprendidas por los defensores de los derechos humanos para ejercer una presión sobre los

procesos reglamentarios y de gobernabilidad de los datos digitales con el objetivo de reducir los riesgos e incrementar su capacidad de acción.

INDEX

Palabras claves : capacidad de actuar, derechos humanos, Estados, gobernabilidad de los datos, prácticas digitales

Keywords : capacity to act, human rights, states, data governance, digital practices

Mots-clés : pratiques numériques, droits humains, États, gouvernance des données, capacité d'agir

AUTEURS

NORMAND LANDRY

Normand Landry est professeur à la TELUQ et titulaire de la Chaire de recherche du Canada en éducation aux médias et droits humains. Courriel : Normand.landry@teluq.ca

JULIEN HOCINE

Julien Hocine est doctorant en communication à l'Université du Québec à Montréal et membre de la Chaire de recherche du Canada en éducation aux médias et droits humains. Courriel : Hocine.julien@courrier.uqam.ca

SIMON FITZBAY

Simon Fitzbay est membre de la Chaire de recherche du Canada en éducation aux médias et droits humains et étudiant au doctorat en communication à l'Université du Québec à Montréal. Courriel : Simon.Fitzbay@teluq.ca

MARIE-JOSÉE DUPUIS

Marie-Josée Dupuis est professionnelle de recherche à la Chaire de recherche du Canada en éducation aux médias et droits humains. Courriel : Marie-Josée.Dupuis@teluq.ca

Vie de laboratoire urbain et quartiers intelligents

Analyse des trajectoires de Sidewalk Toronto et du Quartier des spectacles de Montréal

Etienne Grenier et Nicolas Chartier-Edwards

- 1 Les études critiques consacrées aux enjeux croisés de la surveillance et des infrastructures dédiées à la captation de données numériques (Amoore et Raley, 2017 ; Lyon, 2007 ; Van Dijck, 2014 ; Zuboff, 2020) ont démontré la croissance de ces pratiques dans l'administration urbaine, notamment à travers le développement de quartiers s'inscrivant dans le paradigme de la ville intelligente. Tout en revisitant ces projets ainsi que les technologies automatisées qui y sont mobilisées, mais aussi contestées et détournées, il apparaît pertinent de s'intéresser aux modalités de collecte, de traitement et d'usage des données numériques en tant qu'épicentres de potentielles controverses sociotechniques, révélant par le fait même leur caractère éminemment politique (Callon, 2006). Pour paraphraser la célèbre formule de Langdon Winner (1980) — les artefacts sont-ils politiques ? —, il conviendra alors, comme l'a récemment souligné Kean Birch (2025a) d'évaluer si les données numériques ainsi que les infrastructures permettant leur captation, leur traitement et leur stockage sont également politiques.
- 2 Le présent article traite du caractère fondamentalement ambigu des données numériques produites dans le contexte de zones d'innovations urbaines à partir de deux cas canadiens ainsi que des différentes controverses qu'elles peuvent ou non engendrer : le défunt projet de quartier intelligent Sidewalk Toronto, piloté par la filiale en urbanisme de la maison-mère de Google, Alphabet, ainsi que le déploiement massif de dispositifs de captation vidéo dans le Quartier des spectacles (QdS) à Montréal, sous la bannière de l'initiative du Laboratoire numérique urbain (LNU) et du Projet initial d'infrastructure urbaine numérique (PIIUN) (Normandin, 2018 ; Partenariat du Quartier des spectacles, 2018 ; Ville de Montréal, 2018a). Plus précisément, nous nous pencherons sur les conditions permettant la mise en place et l'opération d'infrastructures de surveillance agissant de façon insidieuse en contexte urbain.

- 3 Plusieurs éléments favorisent la comparaison de ces deux cas : il s'agit de deux projets de développement urbain s'inscrivant dans la mouvance internationale des villes intelligentes (Halpern, 2015 ; Halpern et Mitchell, 2022). Dans les deux cas, une volonté chez les gouvernements municipaux d'accélérer le développement technologique des grandes villes, de renforcer l'économie locale et de présenter une vitrine technologique s'adressant à un public international a encouragé la création de ces quartiers intelligents (Belanger, 2015 ; Bélanger et Lapointe, 2021). De plus, dans ces deux cas, nous avons pu assister à la constitution de partenariats à géométrie variable impliquant des acteurs étatiques, des organisations de la société civile ainsi que des industries (Barns, 2020a). C'est dans l'articulation complexe de la gouvernance de ces partenariats que se sont dessinés les destins opposés des infrastructures de captation et de traitement de données dans le cas du QdS et dans celui du Sidewalk Toronto. Le caractère nébuleux des rapports entourant la propriété des données numériques, les règles encadrant l'opération des infrastructures de surveillance et la nature du consentement citoyen aura permis la mise en place des projets mentionnés ci-dessus sans pour autant garantir leur survie. La comparaison de ces deux cas nous renseigne donc sur la manière dont le caractère politique des données, au sens de la façon dont leur modalité de gestion est débattue ou non dans le cadre d'une controverse, peut orienter le devenir d'un projet d'infrastructure urbaine justement dédiée à la production par captation de ces dernières.

Cadre théorique et méthode

- 4 Afin de produire une étude comparative de ces cas, nous mobilisons et croisons les travaux issus des études des sciences et technologies (STS) (Callon, *op. cit.* ; Venturini, 2010), des études critiques des données (Gillespie, 2014 ; Kitchin, 2014a) et des études critiques des infrastructures (Bowker et Star, 1999 ; Karasti *et al.*, 2016), principalement dans leur déclinaison s'intéressant au paradigme de la ville intelligente (Neumann et Star, 1996 ; Star, 1999).X
- 5 Plus précisément, la méthode de l'analyse des controverses nous aura guidés dans l'identification des deux cas sélectionnés, comparables en ce qu'ils reposaient sur des projets impliquant le déploiement de systèmes de captation de données furtifs en milieu urbain, mais opposés quant à l'évolution des potentielles controverses qui leur seraient associées. En effet, Michel Callon définit les controverses techniques comme « des lieux de négociations » (*op. cit.*, p. 137) où les sociologues peuvent obtenir un état des lieux sur la mise en place des agencements sociotechniques. Ces négociations peuvent évoluer vers des formes de désaccords plus prononcés entre les différents acteurs concernés par le projet. Ce sont justement ces désaccords qui forcent les acteurs à justifier leurs positions sur la place publique, ce qui offre au sociologue des matériaux supplémentaires d'analyse, suivant l'approche de Venturini :
- The notion of disagreement is to be taken in the widest sense: controversies begin when actors discover that they cannot ignore each other and controversies end when actors manage to work out a solid compromise to live together. Anything between these two extremes can be called a controversy (*op. cit.*, p. 261).
- 6 Cette approche méthodologique nous a permis d'identifier les deux cas retenus à des fins comparatives, justement en raison de leur degré différent de problématisation par les divers groupes rassemblés autour des questions de gestion de données numériques,

Sidewalk Toronto ayant été un projet fortement contesté, alors que le cas du QdS a été marqué par l'indifférence. Nous serions respectivement en présence d'une controverse ayant eu une phase d'émergence ainsi qu'une fermeture, au sens classique du concept comme défini par Venturini (*op. cit.*), et d'une controverse « froide », c'est-à-dire en situation paradoxale de « non-controversialité », celle-ci n'ayant pas atteint le stade d'un « problème public largement débattu » (Dandurand, McKelvey et Roberge, 2023 ; McKelvey *et al.*, 2024). Le cas du QdS à Montréal, même s'il comporte la plupart des caractéristiques propres aux controverses associées aux villes intelligentes (dispositif de captation audiovisuelle, partage de données avec des entités privées, redéfinition du rapport des citoyens à la ville, appels à l'éthique purement performatifs) (Halpern et Günel, 2017 ; Hollands, 2015 ; Kitchin, 2014b), a en effet été marqué par une relative absence de débats, de contestations, mis à part une brève couverture journalistique visant le déploiement d'une flotte de caméras.

- 7 La pertinence de la comparaison apparaît à ce point précis. Si le projet Sidewalk Toronto s'est résolu par le retrait de son principal promoteur et son abandon à la suite d'un ensemble d'événements ayant contribué à une escalade de tensions entre la population et l'élite politique torontoise, la fin du projet montréalais se produit en silence : toutes les mentions du LNU ont disparu des communications du Partenariat du Quartier des spectacles (PQdS) et des procès-verbaux du conseil municipal de la Ville de Montréal ; le PIIUN est arrivé à son terme ; un ensemble de caméras mises en réseau couvre une portion du centre-ville de Montréal, opérant silencieusement, à la manière d'un système de surveillance. Cette différence majeure entre les deux cas commande une approche méthodologique différenciée.
- 8 Le projet Sidewalk Toronto n'a jamais vu une seule pelletée de terre avant son abandon. Une ethnographie infrastructurale (Star, *op. cit.*) s'avère impossible. Par conséquent, l'étude de ce cas se concentre sur une analyse médiatique de la controverse portant principalement sur les éléments qui ont contribué à intensifier le débat à propos du projet et à fragiliser la position de Sidewalk Labs à titre de promoteur. Ces événements ont notamment à voir avec des manœuvres effectuées par l'agence à l'insu de son partenaire institutionnel, Waterfront Toronto, pour accaparer le pouvoir et les ressources dans la gestion de projet. Afin de bien circonscrire la controverse dans la lignée des études sur les villes intelligentes, nous nous sommes donc intéressés aux documents associés aux fondements du projet ainsi qu'à la démarche d'origine d'Alphabet, qui ont guidé les processus de conception de Sidewalk Toronto. En raison de ces enjeux, il faut insister sur le fait que les matériaux principaux de l'analyse relèvent de la couverture médiatique, bien étoffée, du déroulement de la controverse.
- 9 Au contraire, le cas montréalais exige de porter une attention à une variété de documents produits par la Ville de Montréal ou un de ses mandataires, le PQdS, afin d'en dessiner le contour. Bien que l'annonce d'un déploiement massif de caméras de surveillance sur le territoire du QdS ait attiré suffisamment l'attention pour amener un journaliste à rédiger un article, nous sommes ici dans une situation diamétralement opposée à celle de Sidewalk Toronto sur le plan de la couverture médiatique. De plus, alors que la littérature grise associée à ce projet est difficilement accessible, voire quasi inexistante, les différents documents produits par le PQdS et la Ville de Montréal dans lesquels on trouve la trace de l'infrastructure numérique de surveillance étudiée (procès-verbaux du conseil municipal, rapports d'activités, fiches descriptives incluses dans les programmes d'immobilisation, brochures promotionnelles) sont légion¹. Cette

donne permet de mener une démarche rappelant l'ethnographie infrastructurelle de Susan Leigh Star, où, grâce à une analyse documentaire, il est possible de suivre à la trace l'émergence d'un système de caméras de surveillance à grande échelle et de remarquer ses différentes incarnations dans les discours institutionnels le concernant (Fairclough, 2012). C'est par le croisement de ces documents que peut être tissé le métarécit de cette infrastructure.

Études de cas

Sidewalk Toronto

Nature du partenariat

- 10 La logique du projet Sidewalk Toronto doit être comprise comme une tentative de conjuguer les ambitions de Sidewalk Labs, une filiale d'Alphabet, la maison-mère de Google, de « créer une cité de l'Internet » grâce à un ensemble de technologies dites intelligentes, aux besoins de la Ville de Toronto désirant revitaliser le district du Quayside, en bordure du lac Ontario. Sidewalk Labs a pour origine un projet étudiant de Sergey Brin et Larry Page, les cofondateurs de Google, ayant pour objectif de réinventer le système de transit routier de la région de San Francisco. Cette impulsion initiale fut traduite en un nouveau projet urbanistique en gestation chez Javelin, le *think tank* d'Alphabet. Ce *think tank* avait pour fonction de procéder à l'idéation purement spéculative (pratique nommée le *moonshot*) de nouvelles technologies devant provoquer une disruption au sein de marchés stables afin d'enclencher de nouveaux cycles d'innovation (Geiger, 2020). Javelin finit par s'attaquer au marché de la planification urbaine à travers la conception d'infrastructures numériques novatrices, visant la construction d'un espace urbain « en circuit fermé », en ce qui a trait aussi bien à sa localisation qu'aux technologies qui y sont mobilisées et sa gouvernance : « What if Javelin built a city at sea, free from any city or country's regulations and restrictions? » (O'Kane, 2022, p. 13). Toutefois, à la suite de tentatives d'acquisitions infructueuses d'un espace géographique libre de régulations, le projet fut révisé afin de l'adapter à une stratégie relevant plutôt de l'insertion, ciblant la ville de Toronto. Sidewalk Labs a donc préféré soumettre une proposition au Waterfront Toronto, une agence publique tripartite établie par les paliers de gouvernement fédéral, provincial et municipal, dans le cadre d'un appel d'offres sous forme de concours visant l'octroi d'un contrat de revitalisation du secteur (Flynn et Valverde, 2019). À ce propos, des journalistes ont rapporté l'existence de rumeurs quant à la nature supposément « viciée » de l'appel d'offres. Ces rumeurs furent confirmées par le dévoilement de l'information suivante : le premier ministre du Canada, Justin Trudeau, et Eric Schmidt, de Google, auraient eu une conversation téléphonique le 16 janvier 2017, soit deux mois avant le déclenchement de l'appel d'offres. Pis encore, cet appel aurait permis de relayer directement l'intérêt d'Alphabet à conclure un partenariat avec Waterfront Toronto afin de construire une ville intelligente (O'Kane, *op. cit.*).
- 11 Le gain net pour Sidewalk Labs d'adopter une stratégie d'insertion est double. D'abord, la compagnie évite de s'embarrasser de la construction d'un cadre bâti *ex nihilo*, pouvant ainsi se concentrer plus directement sur le développement des technologies qui composeront l'infrastructure numérique du quartier. Ensuite, sur le plan du discours promotionnel, la nature même de ce partenariat permet à Sidewalk Labs de se

détacher de l'image monopolistique que projette sa maison-mère ainsi que des projets de villes intelligentes, historiquement catastrophiques, tels que Songdo City et Masdar (Halpern et Günel, *op. cit.* ; Hollands, *op. cit.*). Fondamentalement, cette approche de type public-privé impliquant Sidewalk Labs et Waterfront Toronto, typique du modèle de gouvernance du Parti libéral du Canada en ce qui concerne les industries de l'innovation technologique canadiennes et américaines (Roberge *et al.*, 2019, 2022 ; Senneville-Robert, 2021 ; Smardon, 2020), est à la source des discours promotionnels derrière ce nouveau projet de « zone d'innovation urbaine ».

Participation citoyenne et usage des données

- 12 Historiquement, les discours promotionnels accompagnant les projets de villes intelligentes tendent à mobiliser des ensembles de promesses utopiques, allant de la facilitation de la consommation hebdomadaire de produits jusqu'à la résolution de la crise climatique (Halpern et Günel, *op. cit.*). Le projet Sidewalk Toronto ne fait pas exception. Les désirs de Google, formalisés par Javelin puis Sidewalk Labs dans un document préliminaire intitulé *The Yellow Book*, nous renseignent fortement sur les prémisses du projet :

The Yellow Book declared that Sidewalk would help cities « overcome cynism about the future » and build « a city from the internet up. » The plans proposed a massive community that could house 100,000 people across 1,000 acres. It viewed people through the lenses of efficiency and profit, even as it promised residents better lives through novel technologies. [...] The Yellow Book's back flap imagined a world with « no glass ceiling. No fossil fuels. No taking out the trash. No forgotten passwords » (O'Kane, *op. cit.*, p. 47-48).

- 13 Cette « vie meilleure », garantie par les technologies « intelligentes » ainsi que la capacité de Sidewalk Labs à livrer ces dernières, reste conditionnelle à une forme de participation citoyenne dont les modalités sont entièrement décidées par Sidewalk Labs : le partage des données numériques produites par les citoyens du futur quartier.
- 14 En effet, le projet de ville intelligente imaginé dans le *Yellow Book* de Sidewalk Labs conceptualise un modèle de gouvernance urbaine basée sur une dynamique transactionnelle où la prestation et l'optimisation des services par l'administration de la ville sont conditionnelles à ce que les citoyens fournissent un accès étendu à leurs données numériques privées. La justification offerte par Sidewalk Labs à propos de cet accès est le caractère *indispensable* de ces données numériques, non seulement pour le bon fonctionnement et l'optimisation des infrastructures du quartier, mais, plus encore, pour la recherche et le développement de technologies futures. Cette idée forte, au cœur du modèle de gouvernance de Sidewalk Labs, remonte, elle aussi, aux premiers concepts issus du *Yellow Book* :

Data would be a kind of currency. People who chose not to share anything about themselves when they visited a friend's apartment in Project Sidewalk might not get access to its self-driving taxibots or be able to buy items from certain stores. [...] Data could ultimately be used to reward "good behaviour" (*ibid.*, p. 49-50).

- 15 Cette volonté d'utiliser les données numériques captées dans les espaces publics de Sidewalk Toronto fut centrale dans la réponse de la compagnie à l'appel d'offres, sans toutefois fournir de cadre clair expliquant la gestion de ces dernières (Scassa, 2020). En effet, il importe de contextualiser ce projet dans un certain tournant entrepris en matière de modèle d'affaire de la part de Google et Alphabet autour des années 2010, soit vers la R&D en intelligence artificielle, tournant qui se traduit par l'acquisition de

jeunes pousses comme Deep Mind ainsi que par l'embauche de scientifiques renommés dans le domaine, notamment Geoffrey Hinton, « père fondateur » de l'apprentissage profond (Cardon, Cointet et Mazières, 2018). La possibilité de capter en temps réel un vaste ensemble de données numériques générées par les activités, publiques, mais aussi privées, de la population du futur quartier intelligent présente une occasion d'affaires non négligeable pour une compagnie comme Alphabet. Cette occasion de captation est, d'abord, directement alignée sur les pratiques de traitement de données de Google, qui cherche à découvrir de nouvelles corrélations dans les ensembles de données récoltés grâce à son moteur de recherche (Pasquinelli, 2009 ; Zuboff, *op. cit.*). Les données numériques ne seraient donc pas simplement une « monnaie d'échange » contre des services municipaux, comme présenté dans le *Yellow Book*, elles seraient plus directement la *condition de possibilité même* d'un projet de ville intelligente comme Sidewalk Labs l'envisageait.

- 16 Cette dynamique, qualifiée d'extraction par plusieurs militants, chercheurs universitaires et acteurs de l'industrie canadienne de l'innovation (Artyushina, 2020 ; O'Kane, *op. cit.* ; Senneville-Robert, *op. cit.*), fut plutôt présentée par la direction de Sidewalk Labs comme une forme de participation citoyenne à deux niveaux. Premièrement, le partage des données est envisagé comme une forme de participation citoyenne permettant aux habitants du futur quartier d'effectuer des rétroactions pour contribuer à l'amélioration des services. Cette idée de « plateformes de gouvernance » n'est en rien nouvelle, mais bel et bien constitutive des projets de villes intelligentes (Barns, 2020b ; Hollands, *op. cit.* ; Kitchin, 2014b). Deuxièmement, le partage des données numériques avec Sidewalk Labs devait permettre à cette dernière de remplir l'un des mandats auxquels elle s'était engagée dans le cadre de l'appel d'offres, soit celui de développer et de tester en temps réel de nouvelles technologies sur les futurs citoyens, afin de breveter ces dernières et de partager les profits issus de la commercialisation avec la Ville de Toronto (O'Kane, *op. cit.*). Justement, cette idée du partage de données numériques avec la Ville, présentée par Sidewalk Labs comme un engagement citoyen, aura fait controverse dès les premiers instants de vie du projet, qualifié de toutes parts par un ensemble de détracteurs comme une privatisation des services publics (*ibid.*, p. 123). Pis encore, les exercices de consultation publique menés par Sidewalk Labs dans les mois suivant le démarrage du projet auront plutôt été conçus comme autant de tentatives d'embrigader une large coalition torontoise issue des milieux culturels et de l'innovation afin d'appuyer publiquement le projet comme formulé dans son éventuel plan maître (O'Kane, *op. cit.* ; Senneville-Robert, *op. cit.*). Ces tentatives répétées de Sidewalk Labs d'accaparer la totalité de la gouvernance du projet marquera l'ensemble des manœuvres de la compagnie, jusqu'à l'arrêt et à l'annulation de Sidewalk Toronto.

Problématiques et controverses

- 17 La littérature scientifique ainsi que la couverture journalistique de Sidewalk Toronto insistent sur le moment charnière qu'a constitué le dévoilement du plan de gestion de données dans le développement de la controverse qui torpillera éventuellement le projet (Artyushina, *op. cit.* ; Haggart, 2020 ; O'Kane, *op. cit.* ; Scassa, *op. cit.* ; Senneville-Robert, *op. cit.*). Tout d'abord, il est important de préciser que le gouvernement canadien n'avait pas à l'époque, et n'a toujours pas, de politique publique légiférant sur la gouvernance des données numériques et des technologies d'intelligence artificielle².

Il n'existait donc aucune balise gouvernementale pouvant orienter le plan élaboré par Sidewalk Labs. Il faut également rappeler que, dans les balbutiements du projet, Sidewalk Labs s'était engagée à suivre les principes du *privacy by design*³, proposés par Ann Cavoukian, qui deviendra elle-même consultante pour Sidewalk Labs (O'Kane, *op. cit.* ; Senneville-Robert, *op. cit.*). De plus, le partage des données numériques devait alimenter le développement de nouvelles technologies, ce qui amènerait potentiellement la Ville de Toronto à récupérer une partie des revenus générés par l'aventure. Pourtant, malgré plusieurs conférences de presse, les détails de ce partenariat sont restés largement opaques (O'Kane, *op. cit.*). Ce contexte général, marqué à la fois par le scandale de Cambridge Analytica, par les promesses d'engagement éthique et de retombées économiques ainsi que par des modalités partenariales occultes, a contribué à envenimer a priori la réceptivité des publics face à Sidewalk Toronto.

- 18 La première itération d'un plan structurant du partenariat et des modalités de partage, un document intitulé *Plan Development Agreement*, fut révélée le 31 juillet 2018 à la suite de la démission de deux membres du comité de Waterfront Toronto. Ce document ne contenait toujours que très peu de détails à propos des pratiques de captation de données et de la sécurisation de ces dernières :

The new agreement didn't have much to say about data, for one thing. Beyond a few progressive-but-vague principles and promises to create « citizen-centered » policies recognizing privacy as a human right, and similarly vague assurances that the project would not « foster monopolies » around data or tech, there was still little information about what kind of data would be collected if Quayside were built — or what Sidewalk might do with it (O'Kane, *op. cit.*, p. 160).

- 19 Il n'y avait donc rien dans ce document qui permettait de baliser les pratiques de captation et de traitement des données par Sidewalk Labs. Rien non plus pour réaffirmer le contrôle de l'agence Waterfront Toronto sur le déroulement du projet. Le document n'offrait aucun détail supplémentaire sur la nature de la structure partenariale non plus que sur le partage des profits issus de la commercialisation des technologies brevetées (O'Kane, *op. cit.*, p. 161). À la suite de la parution de ce document, le journaliste du *Globe and Mail* Josh O'Kane découvrit un document confidentiel produit par Sidewalk Labs. Il s'agissait d'un appel à projets s'adressant à des architectes et concepteurs dans lequel Sidewalk Labs se prévalait des droits de propriété sur le type d'habitation modulaire qui serait développé : « All rights, including in the intellectual property, will vest in Sidewalk Labs and Sidewalk Labs will have the opportunity to incorporate it into future phases and other works » (*ibid.*, p. 162). La publication subséquente d'un article du *Globe and Mail* par ce même journaliste révélant la manœuvre de Sidewalk Labs contre son partenaire public (O'Kane, 2018) ne fit qu'aggraver l'état de mécontentement du conseil d'administration de Waterfront Toronto et des Torontois. Cette tentative d'appropriation des brevets allait directement à l'encontre du récit d'origine racontant que les citoyens bénéficieraient directement du partage de données.

- 20 Le plan de gestion de données de Sidewalk Labs s'articulait principalement autour de deux principaux éléments, soit une fiducie nommée le Civic Data Trust et la notion de données urbaines (*urban data*) (Artyushina, *op. cit.* ; O'Kane, 2022 ; Scassa, *op. cit.*). Le Civic Data Trust était défini par Sidewalk Labs comme « an independent entity to control, manage, and make publicly accessible all data that could reasonably be considered a public asset, and a set of rules that would apply to all entities operating in

Quayside, including Sidewalk Lab » (Sidewalk Labs, 2018, p. 10). Cette structure fiduciaire devait notamment rassurer les partenaires du Waterfront Toronto et de la société civile, car dans ce contexte le traitement et le stockage des données ne garantiraient pas un avantage monopolistique à Sidewalk Labs. L'indépendance de la fiducie devait en effet permettre l'accès au contenu de cette dernière par les citoyens et des organisations de la société civile (Artyushina, *op. cit.* ; O'Kane, 2022 ; Scassa, *op. cit.*). Cependant, la notion de données urbaines, comme définie dans le plan de gestion de données, venait limiter le contrôle effectué par la fiducie. En établissant une distinction entre les données urbaines et des données dites conventionnelles, Sidewalk Labs « made the case that only urban data can be considered a public asset and enclosed within the framework of the trust, while the latter type should remain a private asset of the company and its partners » (Artyushina, *op. cit.*, p. 8). Cela signifie que des données captées dans des espaces privés, un commerce par exemple, échappaient au contrôle de la fiducie.

- 21 Profitant du vide législatif canadien entourant la gestion des données numériques dans un tel contexte, Sidewalk Labs a mis de l'avant des concepts qui non seulement l'avantageaient, mais aussi ouvraient la porte à la production d'une réglementation complaisante. Traitant la ville comme un « espace-test » (Halpern et Günel, *op. cit.* ; Krivý, 2018), Sidewalk Labs proposait même de modifier les législations municipales, provinciales et fédérales afin de maximiser les occasions de collecte de données :

From the day Waterfront launched the Quayside RFP, the agency had been eager to see what kind of technological innovations a partner like Sidewalk could bring to Toronto. But this wasn't technology, it was policy, and policy was supposed to come from governments (O'Kane, 2022, p. 187).

- 22 L'incapacité de Sidewalk Labs à fournir un cadre de gestion des données s'arrimant au droit canadien et satisfaisant pour l'ensemble des partenaires aura nécessairement été l'un des facteurs les plus aggravants de l'impopularité grandissante du projet.

Fermeture

- 23 En 2020, à la suite du déclenchement de l'état d'urgence sanitaire au Canada, le projet Sidewalk Toronto fut officiellement annulé par Sidewalk Labs (O'Kane, 2022 ; Senneville-Robert, *op. cit.*). En effet, les deux concepts présentés par Sidewalk Labs devant faire office de cadre de gestion de données n'eurent pour effet que d'intensifier le mécontentement et les critiques à l'égard du projet, mais aussi plus précisément à propos de Sidewalk Labs, dont la réputation s'apparentait désormais à celle d'un mauvais partenaire d'affaires. Le plan de gestion de données de Sidewalk Labs fut même problématisé par différents acteurs sociaux (des organismes de défense des droits civiques ainsi que des représentants du secteur de l'industrie technologique canadienne) comme une première étape dans la mise en place d'un régime de surveillance capitaliste (Artyushina, *op. cit.* ; O'Kane, 2022 ; Senneville-Robert, *op. cit.*). Cet angle de problématisation, adopté par un nombre toujours grandissant de figures importantes gravitant autour de la controverse, engendra de fortes mobilisations citoyennes auprès des décideurs afin de s'assurer que Waterfront Toronto, et non Sidewalk Labs, contrôlait le projet Sidewalk Toronto. Les tentatives de Sidewalk Labs de se distancier des pratiques extractives d'Alphabet et de Google seront infructueuses, car la rentabilité d'un tel projet dépend de la captation et du contrôle des données numériques (Halpern et Günel, *op. cit.* ; Hollands, *op. cit.* ; Krivý, *op. cit.*). Ainsi est mort le

projet de Sidewalk Toronto, plongé dans une forte controverse publique où les tensions entre les différents acteurs ne pouvaient être résolues ou encore ignorées, comme elles le seront au QdS.

Le Quartier des spectacles

Déploiement massif de caméras de surveillance au centre-ville de Montréal

- 24 Le cas montréalais de cette étude comparative ne relève pas, à l'instar de son pendant torontois, d'un mégaprojet clairement identifiable, mené par un puissant et unique acteur du monde des affaires tentant de découper une tranche de l'espace urbain à l'aide d'une infrastructure numérique monolithique. Il est plutôt constitué par l'action concertée de différentes officines municipales en partenariat avec des organisations externes, qu'il s'agisse de mandataires impliqués dans le secteur culturel ou de fournisseurs de technologies, dans un cadre défini par le concept de ville intelligente (Barns, 2020b). L'objet tangible sur lequel repose ce cas a été constitué par le cumul de projets infrastructurels numériques autonomes en plusieurs vagues, de 2018 à 2020. Toutefois, ceux-ci sont distribués sur une même portion de territoire, interconnectés sur un même réseau et agissant selon les mêmes modalités. Il est nécessaire d'évoquer la notion d'assemblage (Akrich, 2006) afin de saisir les contours de cet objet complexe, articulant les caméras de surveillance autour d'un réseau de fibre optique, et de situer l'origine de deux registres de discours différenciés qui lui sont associés, chacun mobilisé pour en justifier le déploiement. L'absence de coordination explicite entre les acteurs portant les différents projets et la mobilisation de ces deux registres de discours compliquent l'analyse de la controverse, tout en contribuant à en expliquer la principale dynamique et son issue. L'acceptabilité sociale concernant le déploiement à grande échelle d'un système de vidéosurveillance sur une portion du centre-ville de Montréal correspondant au QdS est corollaire de la neutralisation du nécessaire débat public qui aurait dû accompagner la mise en place d'une telle infrastructure. La mort dans l'œuf de la controverse potentielle entourant l'installation d'un tel système de données nous oblige à nous tourner vers la composante infrastructurelle matérielle afin d'étudier ce cas. En saisissant les traces laissées par ces objets connectés dans les médias⁴ ainsi que dans la littérature grise produite par la Ville de Montréal⁵ et le PQdS⁶, il devient alors possible de produire le métarécit infrastructurel d'un laboratoire urbain situé au cœur de la cité.
- 25 Jusqu'à récemment, l'utilité d'un déploiement de caméras vidéo à grande échelle était remise en question par certains acteurs municipaux, notamment le Service de police de la Ville de Montréal (SPVM). En 2016, la flotte de caméras de surveillance utilisée par les forces policières se limitait à seulement vingt appareils. Le chef du poste de quartier 21, Simon Durocher, interrogé à l'époque, déclarait que « [l]e besoin ne s'en fait pas sentir. Il n'y a pas un crime en particulier qui justifie l'installation d'une caméra » (Gerbet, 2016). La surveillance vidéo est à l'époque perçue comme étant peu utile par le SPVM, car elle tend à déplacer les problèmes, les criminels n'ayant qu'à éviter la zone couverte par les caméras. De plus, le SPVM s'est fait par le passé réprimander par la Commission d'accès à l'information du Québec, celle-ci soutenant que le service de police ne respectait pas les règles relatives à l'utilisation de ces appareils et qu'il n'était pas en mesure de justifier leur emploi (*ibid.*).

- 26 À l'origine rapportée dans *La Presse*, l'annonce d'un déploiement massif de caméras en 2018 constitue le principal point d'entrée de cette controverse avortée concernant la surveillance vidéo sur le territoire du QdS (Normandin, *op. cit.*). Intitulé « Des caméras seront installées pour évaluer les foules », l'article cite les propos tenus par Audrey Gauthier, porte-parole de la Ville de Montréal. Selon cette dernière, un ensemble de 315 caméras⁷ allait être déployé sur le territoire du QdS afin de procéder au comptage automatique des badauds qui s'y trouvent. L'article mentionne que « ce système permet aussi de déterminer la tendance des mouvements de foule qui est pertinente à l'amélioration des services aux citoyens ainsi que des mesures de sécurité lors des événements grand public » (*ibid.*). Au-delà des simples opérations de dénombrement, il semblerait que ce système de surveillance vidéo soit utilisé par des tiers à des fins de recherche. « Le Quartier des spectacles servira de laboratoire pour un projet de la Ville de Montréal et de l'Université McGill⁸ sur la gestion de foule », écrit le journaliste, dès le début de l'article. La récente acquisition d'un grand nombre de caméras démontre une volte-face de la part du service de police et de la Ville de Montréal. En plus des 315 caméras mentionnées dans l'article, un lot supplémentaire de 415 caméras devait faire l'objet d'une acquisition entièrement consacrée aux activités du SPVM (Ville de Montréal, 2018a)⁹. Le parc de caméras de surveillance municipale serait donc passé de 20 unités à au moins 750 unités en l'espace de deux ans. Ce changement d'attitude peut s'expliquer par une nouvelle donne technologique. L'introduction récente de technologies de l'intelligence artificielle dans le domaine de la surveillance rend maintenant possible le traitement à grande échelle de données visuelles, ce qui facilite la gestion d'un parc étendu d'équipements. Le journaliste James Vincent résume ainsi cette transformation du monde de la surveillance : « Artificial intelligence is giving surveillance cameras digital brains to match their eyes, letting them analyze live video with no humans necessary » (2018). Par exemple, l'automatisation du traitement des données vidéo permet d'extraire automatiquement de l'information à partir d'un ensemble de mégadonnées qu'un humain n'aurait jamais été capable d'analyser. Ce type de fonctionnalité permet de déployer des systèmes de surveillance à grande échelle où l'ensemble d'un territoire est observé sans qu'il y ait de zone occultée.
- 27 Afin d'évaluer la couverture et les usages potentiels d'un tel système, nous avons consulté un ouvrage publié en 2018, *Handbook for Smart Cities*, dans lequel un chapitre écrit par une équipe de chercheurs de McGill dirigée par le professeur Tho Le-Ngoc¹⁰ présente une étude de cas visant le déploiement d'une infrastructure de surveillance vidéo sur le territoire du QdS. Intitulé « Management of video surveillance for smart cities », ce chapitre propose aux lecteurs un survol des principales caractéristiques des systèmes technologiques affectés à la surveillance vidéo dans les villes intelligentes. En guise de prélude à l'analyse d'un potentiel déploiement de caméras de surveillance dans le QdS, les chercheurs soulignent que la plateforme physique de ces appareils les dote de capacités d'analyse telles que la détection et le suivi de mouvements ainsi que la détection de visage (Dao *et al.*, 2018). L'étude de cas a été développée autour de trois scénarios de déploiement et mobilise le concept de couverture vidéo continue. Les chercheurs décrivent ainsi le scénario 2 correspondant approximativement au territoire du QdS :
- Next, we consider the larger commercial district area surrounding Quartier des Spectacles with a similar urban planning attributes (e.g., building types, traffic intersection density, etc.). Scenario 2 covers the north-eastern 1.6 km² half of the

commercial district. Cameras are distributed following the high-density pattern, where a camera is installed at almost every intersection (*ibid.*, p. 303).

- 28 L'équipe estime que le déploiement de 161 caméras sur environ 1,6 km² correspond à un déploiement à haute densité, ce qui signifie une couverture quasi continue de l'espace visé. Sachant que les 315 caméras du PIIUN sont distribuées sur trois sites, il est possible que le QdS en ait reçu la quantité nécessaire afin d'obtenir une telle qualité de couverture.

Métarécit d'une infrastructure urbaine numérique

- 29 Les composantes physiques essentielles de l'infrastructure de surveillance numérique décrites ci-dessus, soit les caméras, le réseau de fibre optique qui les relie et les serveurs informatiques dédiés au traitement et au stockage des flux vidéo, permettent de fixer partiellement l'objet de la quasi-controverse entourant leur déploiement. Cependant, il importe d'intégrer une composante sociale à ces objets connectés afin de pouvoir qualifier pleinement d'*assemblage sociotechnique* l'infrastructure numérique urbaine de surveillance couvrant le territoire du QdS. Composé d'une variété d'entités hétérogènes situées sur différents plans (institutions, machines, humains, etc.), un tel assemblage possède un caractère *émergent* résultant de l'interaction de ces différentes entités (DeLanda, 2013). C'est principalement à travers l'action de deux organisations¹¹, soit le Service des technologies de l'information de la Ville de Montréal et le PQdS, menant chacune un projet structurant et emblématique, respectivement le PIIUN et le LNU, que se constituera progressivement l'assemblage sociotechnique situé au cœur de cette étude de cas. Bien que distincts quant à leur origine organisationnelle, aux objectifs initialement poursuivis et aux discours justifiant leur existence, ces projets se croiseront sur le plan matériel de l'infrastructure numérique urbaine. Il en résultera, à travers l'addition successive d'équipements numériques, une forme de contamination où les propriétés d'un projet glisseront vers l'autre et le mobiliseront afin de répondre à ses propres exigences symboliques. Les discours organisationnels en seront affectés et la situation générale de l'infrastructure gagnera en ambiguïté, c'est-à-dire que l'identité de celle-ci sera *diffuse* tant les acteurs, les justifications et les représentations qui la concernent seront multiples et différenciés. Nous avancerons que cette dernière qualité ayant émergé à travers les actions combinées d'*actants* (Latour, 1994), qu'il s'agisse d'organisations, d'humains ou de machines, permet le maintien et l'opération continue d'une infrastructure de surveillance au cœur de la cité, à l'abri d'une controverse qui aurait pu conduire à une mise en examen plus sérieuse. Afin d'aborder le métarécit de cette infrastructure urbaine numérique, il faudra en fixer les lieux et en décrire les projets de ses principaux acteurs ainsi que les discours qui les soutiennent.

Son territoire : le Quartier des spectacles

- 30 La requalification d'une part du centre-ville de Montréal en espace nommé Quartier des spectacles fut l'un des projets phares de l'administration du maire Tremblay (2002-2012). Des consultations menées auprès d'acteurs culturels associés aux grands événements et aux principaux diffuseurs situés dans cette portion de la cité ont conduit les décideurs à lancer un projet d'envergure qui allait transformer le visage de ce kilomètre carré intra-urbain, soit la constitution du PQdS en 2003. L'obtention en 2007 de l'aval des gouvernements fédéral et provincial (Belanger, *op. cit.*), concrétisé par un budget global de 120 millions de dollars, permettra de lancer plusieurs chantiers et de

créer une programmation culturelle pour le QdS. Maintenant devenu une véritable vitrine touristique participant au *branding* de la Ville de Montréal (Poirier, 2022), le QdS est la nouvelle incarnation ludique d'une section des quartiers centraux. Au-delà de la mise en relation des différents acteurs culturels qui y siégeaient, l'espace urbain fut redessiné, tant sous les coups des pics et des pelles, en témoignent les nouvelles places publiques qui firent leur apparition au fil des ans (Esplanade des arts, Esplanade tranquille, entre autres), que grâce à la mise en place d'un réseau informatisé d'équipements médiatiques destinés à animer les espaces tout en les observant. Ce projet de quartier culturel s'inscrit dans une tendance mondiale visant à mettre en scène la nouvelle économie culturelle et le développement urbain (Belanger, *op. cit.*). La compétition internationale que se livrent les cités afin de prétendre au statut de ville intelligente (Cardullo et Kitchin, 2018), ce qui leur permettrait d'attirer les convoités membres de la classe créative, explique en partie le développement du QdS.

Sa conceptualisation : un laboratoire numérique urbain

- 31 Les villes intelligentes font la promotion de l'innovation technologique (Kitchin, 2014b) et les quartiers culturels situés dans leurs centres ne sont pas en reste. Ils participent à cet élan par le biais de leurs activités et de leurs infrastructures. Abondamment décliné à travers le monde, le projet Sidewalk Toronto en est un exemple ; le concept de laboratoire urbain a permis de justifier le développement de sites d'expérimentations pour un ensemble de technologies reposant sur la captation de mégadonnées urbaines. Le concept de laboratoire urbain implique généralement les éléments suivants : une infrastructure informatique reliée à un ensemble de capteurs couvrant un territoire urbain, des partenariats public-privé permettant l'exploitation des mégadonnées recueillies sur ce territoire ainsi qu'une population plus ou moins avertie responsable d'animer cet assemblage sociotechnique (Hollands, *op. cit.*). Le PQdS propose sa propre déclinaison du laboratoire urbain, le LNU. Dans un communiqué de presse daté de 2015 soulignant la signature d'une entente entre Christie, un manufacturier de projecteurs vidéo, et le PQdS, Jacques Primeau, alors président du conseil d'administration, décrit les paramètres de ce projet. On y mentionne « un potentiel de développement technologique important » ainsi que l'ouverture « de nouvelles perspectives en matière de services publics pouvant transformer le rapport des citoyens à la ville ». Primeau avance même que « le Quartier des spectacles pourrait devenir une rampe de lancement de nouvelles technologies » (PQdS, 2015). Déjà, les grandes lignes du LNU étaient tracées : partenariats public-privé, innovation technologique et altération du lien unissant les usagers d'un quartier à leur propre ville sous l'effet de nouveaux services numériques. C'est plus précisément la nature exacte de ces services et de cette transformation de la citoyenneté qui mérite d'être soulignée dans le récit infrastructurel qui suit.
- 32 D'abord associé au plan lumière¹² du PQdS, le LNU est plus sobrement décrit comme une initiative « favorisant la création d'œuvres d'art numérique ainsi que la recherche et l'innovation au niveau des infrastructures technologiques dans l'espace public » (Ville de Montréal, 2018a, p. 869). À cette époque, selon PQdS, « le réseau de vidéo-projection a un important potentiel de retombées positives » et favorise le rayonnement international de la Ville (*ibid.*). Des ententes de partenariat avec des manufacturiers et les propriétaires des immeubles faisant office de toile permettent le développement de cette composante additionnelle de la scénographie lumineuse du QdS. La mise en

contact des projecteurs vidéo et des appareils d'éclairage à travers le réseau de fibre optique souterrain de la Ville de Montréal facilite la gestion des équipements et leur programmation. Comme le souligne un mémoire déposé par PQdS auprès d'un comité permanent du Patrimoine canadien en 2018, « Le LNU relie les espaces publics du Quartier au moyen d'un réseau de fibres optiques, ainsi que les équipements de vidéoprojections [sic] et plusieurs éléments du mobilier urbain » (PQdS, 2018, p. 6). Le partage de flux médiatique et d'information à travers un réseau de communication destiné à alimenter « installations interactives et projections architecturales » est à l'époque la raison d'être du LNU et de son infrastructure (*ibid.*, p. 7). L'éventuel développement du PIIUN bonifiera cette infrastructure technique par l'ajout de caméras, ce qui contribuera à modifier sa nature et son action.

Sa réalisation matérielle : le Projet initial d'infrastructure urbaine numérique

- 33 En parallèle, le Service des technologies de l'information de la Ville de Montréal s'est progressivement doté à partir de 2018 d'un vaste projet de déploiement d'infrastructures numériques, positionnant encore une fois l'entité municipale comme ville intelligente. Cette initiative s'intitule Projet initial d'infrastructure urbaine numérique (PIIUN) et vise à rendre la Ville maître d'œuvre dans la gestion de son réseau informatique et à connecter à ce dernier un mobilier numérique :

Ce mobilier urbain est utilisé comme plateforme physique d'hébergement d'objets connectés (caméras, sonomètres, capteurs environnementaux, etc.). Ces objets connectés réfèrent au concept de l'internet des objets, un écosystème technologique permettant la captation, le transport, l'analyse et la valorisation des données des objets connectés. L'internet des objets est l'infrastructure de base d'une ville qui souhaite garantir un contexte favorable à son développement (économique, social, etc.) et à son environnement (qualité de vie, sécurité, tourisme, etc.) (Ville de Montréal, 2018a, p. 217).

- 34 Le procès-verbal cité ci-dessus rappelle non seulement l'orientation globale de Montréal dans sa transition vers le statut de ville intelligente, mais contient également les résolutions ayant permis l'acquisition du système de surveillance vidéo mentionné plus tôt. Parmi les objets connectés cités dans le cadre du PIIUN figurent les 315 caméras en partie attribuées au QdS. Celles-ci permettront de « répondre à de multiples besoins de sécurité, de vidéosurveillance et de comptage de foule » (PQdS, 2018, p. 221). Il est important de noter que l'argument de la sécurité est tout de même utilisé pour justifier l'installation du réseau de caméras associé au PIIUN. La lecture du procès-verbal nous permet donc d'affirmer qu'un système de surveillance vidéo a été installé au QdS dans le cadre d'un projet d'infrastructure numérique urbaine typique des villes intelligentes où se côtoient des fonctionnalités propres à la gestion des opérations, à la mobilité urbaine et à la sécurité.
- 35 Cette démarche entreprise par le Service des technologies de l'information de la Ville de Montréal est justifiée par un discours où la cité « cherche à affirmer son virage de ville intelligente et numérique, en déployant de nombreuses solutions innovantes, en proposant et adoptant de nouvelles chartes et politiques technologiques, ou bien encore en s'associant avec des partenaires d'excellence dans le domaine » (Ville de Montréal, 2020, p. 835). Le PIIUN compte parmi ces partenaires Mila¹³, un institut de recherche spécialisé en intelligence artificielle basé à Montréal, afin de profiter de son expertise dans le traitement des données produites par les différents capteurs du projet (caméras et sonomètres). Parmi les objectifs poursuivis par le PIIUN, outre la mise à

l'épreuve d'objets connectés par un déploiement sur le terrain, figurent « l'élaboration d'une gouvernance et une planification d'un programme d'infrastructure urbaine numérique grâce aux leçons apprises sur les différents projets pilotes » (*ibid.*). Toujours selon la Ville de Montréal, l'une des visées de ce futur programme est « d'assurer la protection de la vie privée et l'acceptation sociale ainsi que la sécurité des données » (*ibid.*).

- 36 Il fallut attendre six ans à la suite de l'énonciation du PIIUN afin que soit publié par la Ville de Montréal, en 2024, un document officiel mettant de l'avant un ensemble de 13 principes régissant la captation, le stockage et l'utilisation des données numériques urbaines. Parmi les principes énoncés dans la *Charte des données numériques* se trouve en première position le « droit à la vie privée » accompagné de l'intitulé « Préserver l'intimité et prévenir la surveillance ». La description de ce principe proposée par la Ville de Montréal stipule que celle-ci

s'engage à demeurer vigilante quant au déploiement de toute *technologie comprenant des fonctions permettant l'identification, la localisation ou le profilage*¹⁴ d'une personne, tel que les capteurs ou les caméras. L'usage de ces technologies est donc encadré et les principes de participation publique inclusive sont appliqués dans les réflexions (Ville de Montréal, 2024, p. 8).

- 37 Comme seule mention des caméras dans ce document, il serait donc suggéré que l'encadrement d'une infrastructure capable d'effectuer des tâches de surveillance renvoie aux principes 2 et 12 du même document, respectivement « Inclusion » et « Participation publique ». Si le principe d'inclusion vise principalement à proscrire toute forme d'action discriminatoire potentiellement produite par l'un des maillons de la chaîne technologique dédiée aux données urbaines, le principe de participation publique soutiendrait la mise en place de « *mécanismes consultatifs et délibératifs*¹⁵, favorisant l'inclusion de représentantes et représentants de la société civile, permettant de prendre des décisions collectives autour des données » (*ibid.*, p. 18). Or, le déploiement massif de caméras de surveillance au QdS à partir de 2018 a donné corps à un système de *surveillance*, la Ville elle-même emploie ce terme dans le PIIUN, sans qu'il y ait eu de discussion publique ou de consultations à ce propos.

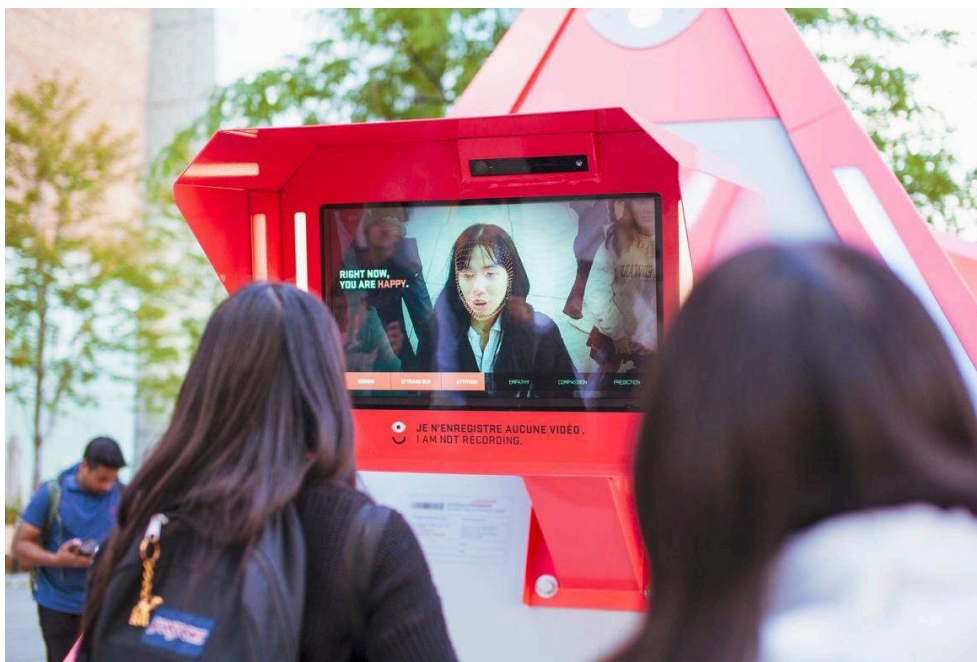
Croisement des récits : l'acceptabilité sociale d'un dispositif de surveillance

- 38 Le dernier principe énoncé dans la *Charte des données numériques*, « Innovation » suivi du sous-titre « Imaginer collectivement grâce aux données », permet de mieux comprendre rétrospectivement ce qui s'est produit lors de l'articulation particulière du réseau de caméras de surveillance au sein du QdS. Le descriptif de principe prend la forme d'un appel à la créativité et à l'innovation ; on y suggère « la possibilité d'imaginer de nouvelles façons d'utiliser les données et les outils numériques » (*ibid.*). De plus, ce principe engage la Ville à « renforcer la capacité d'innovation de la communauté appuyée par les données numériques, en soutenant des espaces d'échange et des initiatives telles que des hackathons et ateliers de cocréation, qui reposent sur l'engagement de parties prenantes autour des enjeux ciblés » (*ibid.*).
- 39 Bien que les *parties prenantes* du descriptif soient les différentes communautés composant la trame sociale de la Ville, il semblerait que celles-ci s'incarnent principalement sous la forme de représentants des industries créatives, par exemple des manufacturiers d'équipements destinés au spectacle tels que Christie ou Panasonic, avec lequel le PQdS a « officialisé une entente de partenariat » en 2021 (PQdS, 2022,

p. 43). Qualifié de « véritable laboratoire urbain, propice à l'expérimentation » (*ibid.*), le QdS est décrit comme un « lieu de démonstration à grande échelle » consacré au « développement technologique et industriel en matière de créativité numérique » plutôt qu'à des exercices de prise de parole citoyenne (PQdS, 2018, p. 1). En offrant le « contrôle des équipements urbains comme l'éclairage de rue ou les caméras de surveillance, l'analyse du déplacement et du trafic au centre-ville, le comptage de foule » (*ibid.*, p. 5), le LNU est une véritable plateforme de tests industriels permettant à différents partenaires corporatifs « de concevoir et mettre en marché de nouveaux produits technologiques » (*ibid.*, p. 6). De plus, on souligne la capacité du LNU à générer « des données qui viendront alimenter le mouvement de données ouvertes dans lequel s'inscrit la Ville de Montréal » (*ibid.*, p. 5), inscrivant ainsi clairement cette dernière dans le paradigme de la ville intelligente.

- 40 Il importe ici de rappeler que les infrastructures technologiques numériques déployées dans les villes intelligentes ont régulièrement été utilisées pour développer des quartiers « ludiques » définis comme « *spaces engendering or allowing a spirit of gaming and playing* » (Vanolo, 2018, p. 322). Le QdS emploie sur une base régulière des capteurs tels que des caméras de surveillance afin de les intégrer dans certaines installations pour leur conférer des propriétés interactives et faire de ce quartier un lieu jouable. Ce mélange des genres s'étend bien au-delà des simples fonctionnalités techniques et opère même sur une couche symbolique, c'est-à-dire qu'il propose une perspective précise sur ce type d'infrastructure, teintée de croyances et de valeurs (Goffman, 2021). Par exemple, *La machine à bienveillance*, une installation interactive coproduite par le collectif Ensemble Ensemble et le PQdS, diffusée en 2019, est ainsi décrite par le site Web du PQdS :

Pour prévenir les comportements déviants, nos sociétés nous scrutent et nous étudient constamment. Mais que se passerait-il si les algorithmes des systèmes de surveillance mesuraient plutôt notre disposition à la bonté ? C'est la question posée ici par Ensemble Ensemble. En se laissant scanner par l'immense caméra, le spectateur reçoit une prédiction personnalisée de ses actes altruistes. Une invitation à refuser la méfiance ordinaire¹⁶ (PQdS, 2022).

Figure 1. *La machine à bienveillance*

Crédits : Photo Chloé Larivière, 2019

- 41 Cette invitation ludique à « refuser la méfiance ordinaire » contribuerait-elle à banaliser la surveillance constante visant les usagers du quartier ? Bien que cette installation n'enregistre aucune vidéo, en ferait foi l'inscription sur son devant, des caméras surveillent l'espace où se trouve son public. Par conséquent, les interactions se déroulant dans le cadre de surveillance ludifiée proposée par cette installation sont fort probablement enregistrées par un second système de surveillance furtif, celui qui surplombe le QdS. Ironiquement, la visibilité de cette installation ludique contribue de facto à l'invisibilisation du second système de vidéosurveillance. La ludification des espaces publics recèle une part d'ombre caractérisée par le guidage des citoyens devenus des interacteurs. Incités à interagir avec différents dispositifs, les citoyens animent un système de captation de données placé en surcouches qui agit bien au-delà des propositions « amusantes » régissant les jeux qui leur sont présentés.
- 42 Une fois le PIIUN complété en 2022¹⁷ et son infrastructure de surveillance déployée, le LNU ne semble plus faire l'objet d'une promotion active par le PQdS. À partir de 2022, aucune mention du Laboratoire n'est faite dans les différents documents produits par le PQdS. Alors que celui-ci était central dans le mandat de l'organisme en 2018, le concept semble définitivement abandonné une fois l'infrastructure déployée. La *Planification stratégique* de 2002, le *Rapport d'activités* ainsi que la *Politique de programmation* de 2023 n'en font aucune mention. La brochure intitulée *Le Laboratoire numérique*, encore disponible en téléchargement sous format PDF sur le site Web du PQdS en décembre 2022, ne l'est plus en 2024. La description des infrastructures numériques correspondant à la matrice technique du LNU peut être identifiée dans le document proposant le renouvellement de la convention de gré à gré unissant PQdS à la Ville de Montréal. Intégré aux procès-verbaux du conseil municipal, ce document mentionne simplement l'existence « d'un réseau de caméras de vidéosurveillance [qui] permet de donner une vue sur le Quartier au SPVM lors de manifestations et aux promoteurs lors

de la présentation de leurs événements » (Ville de Montréal, 2023, p. 1081). Le laboratoire s'est dissipé, mais l'infrastructure de captation de données demeure.

Les données numériques sont-elles politiques ?

- 43 Afin de comprendre l'évolution des cas présentés dans cette étude comparative, nous proposons de revenir à la question posée par Langdon Winner dans son article « Do artifacts have politics? » (1980) en l'orientant vers les données, ainsi que les infrastructures permettant leur extraction et leur traitement. Selon Winner, « the things we call technologies are ways of building order in our world. Many technical devices and systems important in everyday life contain possibilities for many different ways of ordering human activity » (2007, p. 127). Nous proposons d'étendre cette notion aux données numériques elles-mêmes dans la mesure où celles-ci contribuent à l'*ordination* des sociétés contemporaines, comme le soutiennent plusieurs auteurs des études critiques des données numériques et de l'intelligence artificielle (Fourcade et Healy, 2024 ; Halpern et Günel, *op. cit.* ; Kitchin, 2014b ; Krivý, *op. cit.*). Puisque les données numériques produites en contexte urbain le sont dans une perspective d'administration de la vie sociale, elles sont nécessairement associées à des objectifs politico-économiques précis. Par conséquent, elles devraient être l'objet de débats au sein de la société civile tout en agissant en tant que force structurante sur cette dernière. La possibilité même de la mise en débat des données numériques force les promoteurs de ces technologies à en justifier la nécessité. Ces discours déployés par les promoteurs engendrent différents types de réactions dans la société civile, celles-ci étant distribuées sur un spectre allant de la coopération jusqu'à la dissension, en passant par un stade neutre caractérisé par une forme d'indifférence de la société civile. Ces différentes réactions se retrouvent dans les deux cas présentés dans cette étude : du côté de Toronto, malgré une adhésion hâtive de la classe politique, les dissensions citoyennes ainsi que l'incapacité de Sidewalk Labs à justifier la légitimité de son projet auront poussé cette même classe politique à se retirer du projet. À Montréal, l'organisation du projet par une nébuleuse d'acteurs privés et publics a pris la forme d'un cumul progressif de dispositifs technologiques dont les usages sont restés flous, ce qui a contribué à l'indifférence générale le concernant. Si les technologies permettent d'ordonner le monde et que les données numériques existent dans un univers computationnel dont l'ontologie même repose sur l'*ordination de l'information* (Fourcade, 2021 ; Hui, 2016 ; Rieder, 2020), la capacité (ou l'incapacité) des acteurs désirant capter et utiliser ces dernières dépend de la production d'arguments défendant le caractère bénéfique de cette ordination pour l'ensemble de la communauté. Or, la conscientisation progressive du public par rapport aux pratiques de captation de données, qu'elles aient lieu en ligne (Birch *et al.*, 2021 ; McKelvey, 2021 ; Zuboff, *op. cit.*) ou bien en contexte urbain, complique la tâche d'organisations désirant procéder à de telles activités ; l'administration algorithmique des sociétés suppose maintenant que l'on demande à leurs acteurs comment ces derniers veulent « être calculés » (Meunier *et al.*, 2019, p. 203).
- 44 Cette politisation de la question complexifie les modalités de débats. En effet, il apparaît à la lumière des cas étudiés que la production massive de données numériques par l'emploi d'infrastructures de surveillance en contexte urbain ne peut s'exercer que lorsqu'est préservée une forme d'ambiguïté à propos de ce phénomène sociotechnique. Pour répondre à la question de Winner — *do artifacts have politics?* —, nous proposons

que le caractère *politique* des données numériques se révèle lorsque celles-ci et les infrastructures dont elles dépendent alimentent des débats à leur sujet. L'intensité de ces débats devient en elle-même un élément structurant dès lors qu'un projet nécessite une captation de données appartenant à un régime de propriété supposément public. C'est justement la différence d'intensité des débats sur la définition et l'encadrement des données numériques pour Sidewalk Toronto et le LNU qui nous permet de comprendre les trajectoires radicalement différentes de ces deux cas. Leur comparaison révèle que c'est bel et bien la capacité des acteurs sociaux concernés d'effectuer une mise en débat visant ces infrastructures qui décidera de leur sort, ce qui renforce la logique selon laquelle les agencements sociotechniques ne sont pas le pur fruit d'une domination, mais, au contraire, qu'ils sont bel et bien négociables, voire contestables (Akrich, 1989 ; Callon, *op. cit.*). Il apparaît alors que les négociations publiques à propos de ces systèmes de surveillance de masse ont pour effet d'en dévoiler le fonctionnement et de révéler les intérêts de leurs opérateurs et du législateur, ce qui contribue à diminuer leur acceptabilité. Dans le cas torontois, la surpolitisation des données afin de définir les modalités de collecte, de stockage et d'utilisation a provoqué l'arrêt du projet. Dans le cas montréalais, c'est à l'inverse le maintien d'une forme d'ambiguïté entourant la gouvernance des infrastructures de surveillance et la nature des relations unissant une variété d'acteurs étatiques et privés impliqués dans le projet qui en permettra l'opération discrète et continue.

Conclusion

- 45 La comparaison des cas torontois et montréalais nous incite à jeter un regard critique sur les projets infrastructurels considérés comme innovants par les pouvoirs publics et les acteurs industriels. La notion de développement technologique et celle, plutôt vague, de créativité numérique peuvent être mobilisées pour transformer la perception qu'a le public des infrastructures de captation de données en opérant sur une couche symbolique. Les récits construits par les différents acteurs employés dans la mise en place de ces systèmes ainsi que par leur « public », qu'il soit consentant ou non, sont déterminants quant à leur acceptabilité. Ces récits peuvent contribuer au développement d'une controverse sociotechnique autant qu'à sa dilution. De plus, la démultiplication des intervenants et des discours qu'ils produisent afin de motiver un quelconque déploiement infrastructurel ainsi que la production de récits et de représentations différenciés de ce dernier contribuent à complexifier le paysage dans lequel il s'inscrit. L'absence d'un récit « simple », caractérisé entre autres par un nombre limité d'acteurs ayant des motivations explicites et dont les gestes font l'objet d'une évidente coordination, nuit à la constitution d'une éventuelle controverse. Ainsi, la complexité croissante d'une situation contribue au développement d'une ambiguïté qui, in fine, neutralisera de potentiels débats. La production de l'acceptabilité sociale qui en résulte permet la réalisation de ces grands projets qui autrement auraient pu être considérés comme néfastes, au regard de principes élémentaires tels que le droit à la vie privée. Le cas montréalais nous indique même que ce type d'énoncés principaux sont suffisamment vagues et non contraignants qu'ils peuvent être mis de l'avant par une organisation qui agit apparemment en contravention de ceux-ci. De plus, il semblerait que ce soit la quasi-absence de controverse qui permettrait la réussite d'un déploiement d'infrastructure technologique numérique s'inscrivant dans la tradition de la surveillance algorithmique. Ces éléments nous incitent à reconsidérer le cadre

traditionnel de la controverse tel qu'il a été utilisé dans le champ des STS, car il se pourrait qu'il soit de moins en moins adapté à un monde où les rapports sociaux sont redéfinis par le déploiement de dispositifs furtifs destinés à la captation et au traitement de données socionumériques. Nous proposons de revisiter ce cadre conceptuel dans de futurs travaux (Grenier *et coll.*, à paraître) et de nous intéresser précisément aux récits les plus consensuels visant le déploiement de ces infrastructures technologiques.

BIBLIOGRAPHIE

- AKRICH Madeleine (1989), « La construction d'un système socio-technique. Esquisse pour une anthropologie des techniques », *Anthropologie et Sociétés*, 13(2), p. 31-54, DOI : <https://doi.org/10.7202/015076ar>.
- AKRICH Madeleine (2006), « La construction d'un système socio-technique », dans Michel. CALLON et Bruno LATOUR (dir.), *Sociologie de la traduction*, Paris, Presses des Mines, p. 109-134.
- AMOORE Louise et Rita RALEY (2017), « Securing with algorithms: Knowledge, decision, sovereignty », *Security Dialogue*, 48(1), p. 3-10, DOI : <https://doi.org/10.1177/0967010616680753>.
- ARTYUSHINA Anna (2020), « Is civic data governance the key to democratic smart cities? The role of the urban data trust in Sidewalk Toronto », *Telematics and Informatics*, 55, DOI : <https://doi.org/10.1016/j.tele.2020.101456>.
- BARNES Sarah (2020a), « Re-engineering the city: Platform ecosystems and the capture of urban big data », *Frontiers in Sustainable Cities*, 2(32), DOI : <https://doi.org/10.3389/frsc.2020.00032>.
- BELANGER Anouk (2015), « Le projet du Quartier des spectacles de Montréal : Une mise en scène de l'économie culturelle », dans Myrtille ROY-VALEX et Guy BELLAVANCE (dir.), *Art et territoire à l'ère du développement durable. Vers une nouvelle économie culturelle ?*, Québec, Presses de l'Université Laval, https://www.academia.edu/38457488/Le_projet_du_Quartier_des_spectacles_de_Montr%C3%A9al_Mise_en_sc%C3%A8ne_de_l%C3%A9conomie_culturelle_pdf, page consultée le 7 décembre 2025.
- BÉLANGER Hélène et Dominic LAPOINTE (2021), « Revitalisation et “bulles touristiques” : Une gentrification instantanée par la touristification du quotidien ? », *Recherches sociographiques*, 62(1), p. 149-173, DOI : <https://doi.org/10.7202/1082616ar>.
- BIRCH Kean (2025a), « Do artifacts have political economy? », *Science, Technology & Human Values*, août, DOI : <https://doi.org/10.1177/01622439251352167>.
- BIRCH Kean, DT COCHRANE et Callum WARD (2021), « Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech », *Big Data & Society*, 8(1), DOI : <https://doi.org/10.1177/20539517211017308>.
- BOWKER Geoffrey C. et Sarah STARLEIGH (1999), *Sorting Things Out: Classification and its Consequences*, Londres, MIT Press.

CALLON Michel (2006), « Pour une sociologie des controverses technologiques », dans *Sociologie de la traduction : Textes fondateurs*, Paris, Presses des Mines, p. 135-157.

CARDON Dominique, Jean-Philippe COINTET et Antoine MAZIÈRES (2018), « La revanche des neurones. L'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, 211(5), p. 173-220, DOI : <https://doi.org/10.3917/res.211.0173>.

CARDULLO Paolo et Rob KITCHIN (2018), « Smart urbanism and smart citizenship: The neoliberal logic of “citizen- focused” smart cities in Europe », *The programmable City*, 39, <https://kitchin.org/wp-content/uploads/2019/04/PCP-WP-39-smart-citizenship.pdf>, page consultée le 7 décembre 2025.

COMMISSAIRE À L'INFORMATION ET À LA PROTECTION DE LA VIE PRIVÉE DE L'ONTARIO (2018), *Rapport Privacy by Design*.

COMMISSION DE LA SÉCURITÉ PUBLIQUE (2021), *Utilisation par le SPVM de technologies de reconnaissance faciale et de systèmes de reconnaissance de plaques d'immatriculation*, Montréal, Ville de Montréal.

DANDURAND Guillaume, Fenwick MCKELVEY et Jonathan ROBERGE (2023), « Freezing out: Legacy media's shaping of AI as a cold controversy », *Big Data & Society*, 10(2), DOI : <https://doi.org/10.1177/20539517231219242>.

DAO Nhat-Quang, Quang LE-DANG, Robert MORAWSKI, Anh-Tuan DANG et Tho LE-NGOC (2018), « Management of Video Surveillance for Smart Cities », dans Muthucumaru MAHESWARAN et Elarbi BADIDI (dir.), *Handbook of Smart Cities : Software Services and Cyber Infrastructure*, New York, Springer International Publishing, p. 285-310, DOI : https://doi.org/10.1007/978-3-319-97271-8_11.

DELANDA Manuel (2013), *A New Philosophy of Society: Assemblage Theory and Social Complexity*, Londres, Bloomsbury Academic.

FAIRCLOUGH Norman (2012), « Critical discourse analysis », dans *The Routledge Handbook of Discourse Analysis*, Londres, Routledge.

FLYNN Alexandre et Mariana VALVERDE (2019), « Where the sidewalk ends: The governance of Waterfront Toronto's Sidewalk labs deal. Windsor yearbook of access to justice », 36, p. 263-283, DOI : <https://doi.org/10.22329/wyaj.v36i0.6425>

FOURCADE Marion (2021), « Ordinal citizenship », *The British Journal of Sociology*, 72(2), p. 154-173, DOI : <https://doi.org/10.1111/1468-4446.12839>.

FOURCADE Marion et Kieran HEALY (2024), *The Ordinal Society*, Londres, Harvard University Press.

GEIGER Susi (2020), « Silicon Valley, disruption, and the end of uncertainty », *Journal of Cultural Economy*, 13(2), p. 169-184, DOI : <https://doi.org/10.1080/17530350.2019.1684337>.

GERBET Thomas (2016), « La police de Montréal délaisse les caméras de surveillance », *Radio-Canada*, Radio-Canada.ca, 12 mai, <https://ici.radio-canada.ca/nouvelle/780487/cameras-spvm-montreal-surveillance-police>, page consultée le 7 décembre 2025.

GILLESPIE Tarleton (2014), « The relevance of algorithms », Tarleton GILLESPIE, Pablo J. BOCZKOWSKI et Kirsten FOOT (dir.), *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, The MIT Press, p. 167-193.

GOFFMAN Erving. (2021), *The Presentation of Self in Everyday Life*, New York, Knopf Doubleday Publishing Group.

- HAGGART Blayne (2020), « The selling of Toronto's smart city », *Smart Cities in Canada, Digital Dreams, Corporate Designs*, Toronto, James Lorimer & Company, p. 38-51.
- HALPERN Orit (2015), *Beautiful Data: A History of Vision and Reason Since 1945*, Durham, Duke University Press.
- HALPERN Orit et Gökçe GÜNEL (2017), « Demoing unto death : Smart cities, environment, and preemptive hope », *The Fibreculture Journal*, article 215, <https://twenty-nine.fibreculturejournal.org/fcj-215-demoing-unto-death-smart-cities-environment-and-preemptive-hope/>, page consultée le 7 décembre 2025.
- HALPERN Orit et Robert MITCHELL (2022), *The Smartness Mandate*, Londres, The MIT Press.
- HOLLANDS Robert G. (2015), « Critical interventions into the corporate smart city », *Cambridge Journal of Regions, Economy and Society*, 8(1), p. 61-77, DOI : <https://doi.org/10.1093/cjres/rsu011>.
- HUI Yui. (2016), *On the Existence of Digital Objects*, London, University of Minnesota Press.
- KARASTI Helena, Florence MILLERAND, Christine M. HINE et Christine G. BOWKER (2016), « Knowledge infrastructures: Part I », *Science & Technology Studies*, 29(1), p. 2-12, DOI : <https://doi.org/10.23987/sts.55406>.
- KITCHIN Rob (2014a), « Big Data, new epistemologies and paradigm shifts », *Big Data & Society*, 1(1), DOI : <https://doi.org/10.1177/2053951714528481>.
- KITCHIN Rob (2014b), « The real-time city? Big data and smart urbanism », *GeoJournal*, 79(1), p. 1-14, DOI : <https://doi.org/10.1007/s10708-013-9516-8>.
- KRIVÝ Maros (2018), « Towards a critique of cybernetic urbanism: The smart city and the society of control », *Planning Theory*, 17(1), p. 8-30, DOI : <https://doi.org/10.1177/1473095216645631>.
- LATOUR Bruno (1994), « Une sociologie sans objet ? Note théorique sur l'interobjectivité », *Sociologie du travail*, 36(4), p. 587-607, DOI : <https://doi.org/10.3406/sotra.1994.2196>.
- LYON David (2007), *Surveillance Studies: An Overview*, Bodmin, Polity.
- MCKELVEY Fenwick (2021), « The Other Cambridge Analytics: Early "Artificial Intelligence" in American Political Science », dans Jonathan ROBERGE et Michael CASTELLE (dir.), *The Cultural Life of Machine Learning : An Incursion into Critical AI Studies*, Cham, Springer International Publishing, DOI : https://doi.org/10.1007/978-3-030-56286-1_4.
- MCKELVEY Fenwick, Joanna REDDEN, Jonathan ROBERGE et Luke STARK (2024), « (Un)stable diffusions: The publics, publicities, and publicizations of generative AI », *Journal of Digital Social Research*, 6(4), p. 1-12, DOI : <https://doi.org/10.33621/jdsr.v6i440453>.
- MEUNIER Axel, Donato RICCI, Dominique CARDON et Maxime CRÉPEL (2019), « Les glitches, ces moments où les algorithmes tremblent », *Techniques & Culture. Revue semestrielle d'anthropologie des techniques*, DOI : <https://doi.org/10.4000/tc.12594>.
- NEUMANN Laura J. et Susan Leigh STAR (1996), « Making infrastructure: The dream of a common language », dans *Proceedings of the Participatory Design Conference*, 13-15 novembre, Cambridge, Computer Professionals for Social Responsibility, p. 231-240.
- NORMANDIN Pierre-André (2018), « Quartier des spectacles : Des caméras pour évaluer les foules », *La Presse*, 15 octobre, <https://www.lapresse.ca/actualites/grand-montreal/201810/14/01-5200224-quartier-des-spectacles-des-cameras-pour-evaluer-les-foules.php>, page consultée le 7 décembre 2025.
- O'KANE Josh (2018), « Sidewalk Labs taking steps to control intellectual property on Toronto's "smart city" document shows », *The Globe and Mail*, 31 octobre, <https://>

www.theglobeandmail.com/business/article-sidewalk-labs-taking-steps-to-control-intellectual-property-on-toronto/, page consultée le 7 décembre 2025.

O’KANE Josh (2022), *Sideways: The City Google Couldn’t Buy*, Toronto, Random House Canada.

PARTENARIAT DU QUARTIER DES SPECTACLES (PQds) (2015), « Christie Digital s’associe au Laboratoire numérique urbain du Quartier des spectacles », *V2com newswire*, 6 octobre, <https://www.v2com-newswire.com/fr/salle-de-presse/categories/design-d-eclairage/dossiers-de-presse/583-10/christie-digital-s-associe-au-laboratoire-numerique-urbain-du-quartier-des-spectacles>, page consultée le 7 décembre 2025.

PARTENARIAT DU QUARTIER DES SPECTACLES (PQds) (2018), « Laboratoire numérique urbain », Montréal, Partenariat du Quartier des Spectacles, p. 1-9).

PARTENARIAT DU QUARTIER DES SPECTACLES (PQds) (2022), *Rapport d’activités 2021*, Montréal, Partenariat du Quartier des Spectacles.

PASQUINELLI Matteo (2009), « Google’s PageRank algorithm: A diagram of cognitive capitalism and the rentier of the common intellect », *Deep search: The politics of search beyond Google*, p. 152-162.

PICKREN Graham (2018), « The global assemblage of digital flow’: Critical data studies and the infrastructures of computing », *Progress in Human Geography*, 42(2), p. 225-243, DOI : <https://doi.org/10.1177/0309132516673241>.

POIRIER Josianne (2022), *Montréal fantasmagorique : Ou la part d’ombre des animations lumineuses urbaines*, Montréal, Lux Éditeur.

RIEDER Bernhard (2020), *Engines of Order : A Mechanology of Algorithmic Techniques*, Amsterdam, Amsterdam University Press, DOI : <https://doi.org/10.2307/j.ctv12sdvf1>.

ROBERGE Jonathan, Guillaume DANDURAND, Kevin MORIN, Kevin et Marius SENNEVILLE (2022), « Les narvals et les licornes se cachent-ils pour mourir ? De la cybernétique, de la gouvernance et d’Element AI », *Réseaux*, 232-233(2-3), p. 169-196, DOI : <https://doi.org/10.3917/res.232.0169>.

ROBERGE Jonathan, Kevin MORIN et Marius SENNEVILLE (2019), « Deep learning’s governmentality: The other black box », dans Andreas SUDMANN (dir.), *The Democratization of Artificial intelligence*, Bielefeld, transcript Verlag, p. 123-142, DOI : <https://doi.org/10.1515/9783839447192-008>.

SCASSA Theresa (2020), « Designing Data Governance for Data Sharing: Lessons from Sidewalk Toronto », *Technology and Regulation*, 44(56), <https://papers.ssrn.com/abstract=3722204>, page consultée le 7 décembre 2025.

SENNEVILLE Marius (2021), *Reconfiguration des liens de collaboration entre acteurs industriels et universitaires de la recherche en intelligence artificielle à Montréal et à Toronto*, mémoire de maîtrise en études urbaines, Montréal, Institut national de la recherche scientifique.

SIDEWALK LABS (2018), « Digital Governance Proposals for DSAP Consultation », *Sidewalk Labs*, octobre, <https://www.waterfrontoronto.ca/sites/default/files/documents/18-10-16-swt-draft-proposals-regarding-data-use-and-governance-tuesday-730pm.pdf>, page consultée le 7 octobre 2025.

SMARDON Bruce (2020), « Dependent technological change: The long arc of canadian innovation and political economic development », dans Heather WHITESIDE (dir.), *Canadian Political Economy*, Toronto, University of Toronto Press, p. 42-68, DOI : <https://doi.org/10.3138/9781487523480.004>.

- STAR Susan Leigh (1999), « The ethnography of infrastructure », *American behavioral scientist*, 43(3), p. 377-391, DOI : <https://doi.org/10.1177/00027649921955326>.
- VAN DIJCK José (2014), « Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology », *Surveillance & Society*, 12(2), p. 197-208, DOI : <https://doi.org/10.24908/ss.v12i2.4776>.
- VANOLO Alberto (2018), « Cities and the politics of gamification », *Cities*, 74, p. 320-326, DOI : <https://doi.org/10.1016/j.cities.2017.12.021>.
- VENTURINI Tomasso (2010), « Diving in magma: How to explore controversies with actor-network theory », *Public Understanding of Science*, 19(3), p. 258-273, DOI : <https://doi.org/10.1177/0963662509102694>.
- VILLE DE MONTRÉAL (2018a), *Assemblée ordinaire du conseil d'agglomération* (v1), Montréal.
- VILLE DE MONTRÉAL (2018b), *Programme triennal d'immobilisations Montréal*, PTI 2019-2021, Montréal.
- VILLE DE MONTRÉAL (2019), « Fiches détaillées », *Programme triennal d'immobilisations 2020-2022*, Montréal.
- VILLE DE MONTRÉAL (2020), *Assemblée ordinaire du conseil d'agglomération* (v1), Montréal.
- VILLE DE MONTRÉAL (2023), *Assemblée ordinaire du conseil d'agglomération* (v1), Montréal.
- VILLE DE MONTRÉAL (2024), *Charte des données numériques*, septembre, Montréal.
- VINCENT James. (2018), « Artificial intelligence is going to supercharge surveillance », *The Verge*, 23 janvier, <https://www.theverge.com/2018/1/23/16907238/artificial-intelligence-surveillance-cameras-security>, page consultée le 7 décembre 2025.
- WINNER Langdon (1980), « Do artifacts have politics? », *Daedalus*, 109(1), p. 121-136.
- ZUBOFF Shoshana (2020), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (4th éd.), New York, PublicAffairs.

NOTES

1. Les documents ont été récupérés en deux phases de moissonnage manuel distinctes. La première a eu lieu en décembre 2023 et la seconde, en février 2025. Des courriels ont été échangés avec le secrétariat de la Commission de la sécurité publique de la Ville de Montréal ainsi qu'avec le Service du greffe de la Ville de Montréal afin d'obtenir certains documents et des compléments d'information.
2. L'absence d'encadrement des technologies de reconnaissance faciale par le gouvernement fédéral sera rappelée dans la description du cas montréalais, la situation laissant les élus municipaux en plan, contraints d'adopter des résolutions ad hoc pour réduire les problèmes qu'elles pourraient causer (Commission de la sécurité publique, 2021).
3. L'expression *privacy by design* désigne une méthodologie visant à intégrer de manière proactive, dès leur conception, des mesures de protection de la vie privée au sein des infrastructures associées au domaine des technologies de l'information et de la communication (Commissaire à l'information et à la protection de la vie privée de l'Ontario, 2018).
4. La question de la surveillance vidéo à Montréal a été brièvement abordée par Radio-Canada et *La Presse*. Cette couverture médiatique (2016-2018) a précédé un débat concernant l'emploi de technologies de reconnaissance faciale par les forces de l'ordre. À ce titre, un rapport concernant

ces technologies a été déposé par la Commission de la sécurité publique de la Ville de Montréal le 17 juin 2021 dans le sillage de l'affaire Clearview AI (Commission de la sécurité publique, *op. cit.*).

5. Des extraits de procès-verbaux de l'assemblée ordinaire du conseil municipal (Ville de Montréal, 2018a, 2020, 2023) ainsi que des extraits de Programmes triennaux d'immobilisations (PTI) (Ville de Montréal, 2018b, 2019) nous ont permis de retracer les décisions concernant le déploiement des caméras sur le territoire du QdS ainsi que leur incorporation dans le cadre du PIIUN de la Ville de Montréal.

6. Des rapports d'activités, des mémoires, des communiqués de presse et des brochures promotionnelles produits par le PQdS nous ont permis de suivre l'évolution du LNU et son enchevêtrement dans la trame infrastructurelle déployée par la Ville de Montréal, notamment à travers le PIIUN.

7. Cette information n'est pas exacte. Le procès-verbal du conseil d'agglomération ayant entériné l'achat des caméras indique plutôt que le lot de 315 caméras est en fait intégré au PIIUN et qu'il sera réparti sur plusieurs sites, y compris le QdS. Espace pour la vie (Biodôme, Biosphère, Jardin botanique, Insectarium et Planétarium) et le Complexe environnemental de Saint-Michel sont les deux autres lieux où seront distribuées les caméras.

8. L'étude des décisions recueillies dans les procès-verbaux du conseil municipal du 14 décembre 2020 révèle plutôt qu'une entente de gré à gré d'une valeur de 366 195,38 \$ a été conclue avec le Montreal Institute for Learning Algorithms (Mila), afin « d'obtenir de l'expertise de pointe en recherche et développement technologique » concernant le traitement des données captées par les caméras.

9. Un contrat d'une valeur potentielle de 590 489,75 \$ a été accordé à la compagnie Panavideo en 2018 « pour l'acquisition de caméras et d'une solution de vidéosurveillance avec service d'installation et d'intégration » dans le cadre de ce projet (Ville de Montréal, 2018a).

10. Il n'a pas été possible de vérifier si le professeur Tho Le-Ngoc était le chercheur principal venant de l'Université McGill associé au projet municipal annoncé par la porte-parole de la Ville de Montréal.

11. Une kyrielle d'acteurs supplémentaires tels que des manufacturiers, des fournisseurs, des scientifiques, des cadres et des techniciens se grefferont au projet au fil de son évolution.

12. Le plan lumière est une initiative de scénographie lumineuse du QdS impliquant différents équipements d'éclairage ainsi que des projecteurs destinés à animer les façades de bâtiments emblématiques de ce quartier grâce à des techniques de *mapping* vidéo. Par exemple, d'iconiques points rouges dessinés par des projecteurs lumineux permettent de reconnaître les établissements et les salles de spectacles associées au PQdS.

13. Regroupant des chercheurs issus de plusieurs universités, Mila est notamment reconnu pour ses nombreux partenariats industriels ainsi que pour son implication dans un vaste écosystème de jeunes pousses. Voir Mila, *À propos de Mila*, <https://mila.quebec/fr/a-propos/a-propos-de-mila>, page consultée le 11 avril 2025.

14. L'italique provient du texte cité.

15. L'italique provient du texte cité.

16. Site Web du PQdS, consulté en 2022, hors ligne actuellement. <https://www.quartierdesspectacles.com/fr/a-propos/les-productions-du-partenariat/oeuvre/79/la-machine-a-bienveillance-par-ensemble-ensemble>.

17. Un investissement total de 5 406 000 \$ aura été injecté dans ce projet selon l'information rapportée dans les fiches détaillées du *Programme triennal d'immobilisations 2020-2022* (Ville de Montréal, 2019).

RÉSUMÉS

Cet article examine le caractère controversé des infrastructures de captation audiovisuelle s'inscrivant dans le paradigme de la ville intelligente, au travers de cas situés à Toronto (Sidewalk Toronto) et à Montréal (Laboratoire numérique urbain). C'est à partir de l'approche théorique des études critiques des données que nous démontrons comment le caractère controversé entourant l'utilisation d'infrastructures de captation dans ces contextes oriente leur devenir. La comparaison de ces cas illustre comment l'ambiguïté entourant l'usage de ces données dilue les controverses et neutralise les débats publics.

This article examines the controversial nature of audiovisual capture infrastructure as part of the smart city paradigm, through cases in Toronto (Sidewalk Toronto) and Montréal (Laboratoire numérique urbain). Using the theoretical approach of critical data studies, we show how the controversy surrounding the use of capture infrastructure in these contexts guides their future. A comparison of the cases illustrates how the ambiguity surrounding the use of this data weakens controversy and neutralizes public debate.

El artículo se refiere al análisis del carácter controvertido de las infraestructuras de captación audiovisual enmarcada en el paradigma de la ciudad inteligente, a través de los casos existentes en Toronto (Sidewalk Toronto) y en Montreal (Laboratorio digital urbano). A partir de un enfoque teórico de los estudios detallados de los datos, demostramos de qué manera el carácter controvertido en torno a la utilización de infraestructuras de captación en esos contextos orienta su futuro. La comparación de esos casos expone cómo la ambigüedad relativa al uso de estos datos atenúa las controversias y neutraliza los debates públicos.

INDEX

Palabras claves : controversias, datos digitales, infraestructura, vigilancia, ciudad inteligente

Keywords : controversies, digital data, infrastructure, surveillance, smart city

Mots-clés : controverses, données numériques, ville intelligente, infrastructure, surveillance

AUTEURS

ETIENNE GRENIER

Etienne Grenier est candidat au doctorat à l'Institut national de la recherche scientifique.

Courriel : etienne.grenier@inrs.ca

NICOLAS CHARTIER-EDWARDS

Nicolas Chartier-Edwards est candidat au doctorat à l'Institut national de la recherche scientifique. Courriel : nicolas.chartier-edwards@inrs.ca

Pour l'amour de la donnée

Vers une datafication de nos affects ?

Camille Alloing, Ambre Marionneau, Julien Pierre et Fabien Richert

- 1 « Les messages en ligne influencent notre expérience des émotions, ce qui peut avoir une incidence sur toute une série de comportements hors ligne¹ » (Kramer, Guillory et Hancock, 2014). Cette affirmation issue d'une étude controversée (Metcalf et Crawford, 2016) dirigée par un chercheur de Facebook met en avant la possibilité de générer une « contagion émotionnelle » parmi les utilisateurs de la plateforme de médias sociaux. Cette hypothèse, corroborée par d'autres travaux (Berger et Milkaman, 2013 ; Tellis *et al.*, 2019), s'inscrit dans le tournant affectif des industries du numérique (Alloing et Pierre, 2021a). Stimulée depuis trente ans par le développement de « l'informatique affective » (Picard, 2000), la promesse de « contagion émotionnelle » amène à considérer les plateformes du Web social comme des « infrastructures affectives » (Hallinan, 2019), proposant à leurs publics (Papacharissi, 2015) des « affectives timelines » (Gilroy-Ware, 2017), à travers lesquelles il est possible d'exprimer des émotions et d'être affecté en retour. Le capitalisme de plateforme (Srnicsek, 2017) s'inscrit lui-même dans un capitalisme affectif (Karppi *et al.*, 2016). Les plateformes qui souhaitent s'insérer dans ces logiques capitalistiques développent un ensemble de fonctionnalités pour produire des signes affectifs comme les *hashtags* (Papacharissi et de Fatima Oliveira, 2012) ou les émojis (Stark et Crawford, 2015). Elles permettent de les faire circuler par diverses fonctionnalités de partage et de les attacher à des contenus par un design émotionnel (Norman, 2007) dont les affordances orientent les usages (Sturm Wilkerson, Riedl et Whipple, 2021 ; Alloing et Pierre, 2021b).
- 2 Cependant, le développement de « machines qui ont l'habileté de reconnaître les états affectifs humains » (Picard *et al.*, 2001, p. 1175) nécessite la « collecte de données affectives de qualité » (*ibid.*, p. 1178) autant physiologiques (voix, mouvement, etc.) que sémiotiques (textes, images, etc.). Les plateformes vont ainsi produire ces données affectives afin de définir un sentiment ou d'identifier un ressenti (Poria, Cambria et Gelbukh, 2015 ; Alloing *et al.*, 2025). Elles peuvent ensuite utiliser ces données pour trier, classer et hiérarchiser les informations qu'elles mettent en visibilité, afin par exemple de nourrir leurs « intelligences artificielles ». Selon Arvidsson (2011), cette

capacité à produire des données affectives permet la production d'un « sentiment général » apte à évaluer certaines valeurs intangibles (la réputation, par exemple) d'une entité. Le traitement algorithmique des émotions a ainsi le potentiel de renforcer les actifs informationnels des plateformes, et par relation contractuelle ceux des organisations commerciales ou politiques qui ont accès à ce type de données. Les plateformes ont alors tout intérêt à communiquer sur cette possible production de valeur pour leurs partenaires commerciaux.

- 3 Peu d'études en sciences sociales explorent les discours promotionnels tenus par les plateformes de médias socionumériques à propos de leur usage et de leur représentation des émotions et des affects (Alloing et Pierre, 2017 ; Hillis, Paasonen et Petit, 2015 ; Stark, 2020 ; McStay, 2020). De même, aucune étude ne recense à ce jour les moyens de captation et les formes de production des données émotionnelles employées par les plateformes, et pourtant décrites dans les publications scientifiques des équipes de ces plateformes ou dans les brevets déposés par leurs ingénieurs. Si ces technologies ne semblent pas être au centre de leurs investissements, nous émettons toutefois l'hypothèse qu'elles sont le reflet de certaines représentations que les concepteurs, ingénieurs et promoteurs des plateformes se font des usagers et usagères et des relations qu'il est possible de construire à partir de leurs émotions — ou des données censées les représenter. Analyser les moyens mis en œuvre pour produire des données sur les affects et émotions des usagers et usagères, ainsi que les objectifs associés à leur manipulation et à leur traitement, nous apparaît comme une occasion de porter un regard critique sur les mécanismes de gouvernance de ces plateformes. En somme, nous proposons d'articuler nos propos autour de la question suivante : comment les plateformes disent-elles opérer la *datafication* de nos affects ? Nous souhaitons par là enrichir les travaux portant sur la captation de l'attention par les plateformes ou encore la viralité des contenus par une description fine de la manière dont les émotions sont conceptualisées par les ingénieurs et valorisées par les équipes marketing de ces plateformes.
- 4 Pour cela, nous nous appuyons sur trois corpus distincts : la littérature scientifique produite par les équipes de recherche de Meta et de Google concernant les émotions, les demandes de brevets déposées depuis plus de dix ans par Meta, Google, ByteDance et X/Twitter et mentionnant des traitements émotionnels et les discours concernant l'exploitation des émotions individuelles tels qu'on peut les trouver dans la communication officielle de ces quatre entreprises. À partir de ces analyses, nos principaux résultats sont discutés sous trois angles. Dans le premier, nous présentons les promesses et représentations associées à l'usage des données affectives. Puis, nous analysons comment ces plateformes construisent concrètement des données affectives en captant et en interprétant les données biométriques et informatiques stockées dans leurs bases de données. Enfin, nous proposons une réflexion critique au sein de laquelle le processus de *datafication* est vu comme un vaste mouvement de captation du vécu émotionnel des usagers et usagères entraînant une servitude passionnelle à l'égard de ces plateformes. Nous concluons en affirmant que, pour comprendre la mise en données des émotions humaines par les plateformes numériques, il est utile de repenser la *datafication* comme un ensemble de représentations à l'œuvre parmi les ingénieurs, chercheurs et concepteurs.

Émotions numériques : la *datafication* de l'intime

- 5 La très grande majorité des humains a déjà fait l'expérience d'une émotion, bien avant le numérique. Mais combien sont capables de dire ce que c'est ? Combien même sont capables de dire si le mot choisi convient pour décrire l'expérience vécue ? Même les poètes n'y arrivent pas. Chez les scientifiques, deux perspectives théoriques s'opposent, l'une culturelle et l'autre biologique. En effet, l'émotion reste un objet controversé en sciences. Nous proposons ici de présenter les principales théories en vigueur, pour mieux comprendre pourquoi l'une d'entre elles est privilégiée par les tenants de l'informatique affective, et comment sa supposée scientificité consolide la promesse de performance que les plateformes font aux marques et partis.

Approches distinctes des émotions

- 6 Dans une approche culturelle, les émotions sont des signes, donc ouverts à des potentialités d'articulation et d'interprétation tributaires des conditions sociales. Le deuil, entre autres exemples, ne s'exprime pas de la même manière selon les zones culturelles (Mauss, 1921) ou les périodes historiques (Corbin, Courtine et Vigarello, 2016a, 2016b, 2017). Pour le sociologue Julien Bernard (2015), les émotions peuvent être considérées comme des constructions sociales (depuis Émile Durkheim) ou des indicateurs d'un rang social (chez Pierre Bourdieu), qui peuvent cependant s'inscrire dans des dynamiques non déterministes. Des régimes émotionnels, avec plus ou moins de contrôle pendant l'époque victorienne ou d'extimité à l'heure des médias sociaux, alterneraient ainsi selon les époques. Analyser ces « structures de sentiments » (selon l'expression de Raymond Williams) permettrait ainsi de mieux saisir les valeurs propres à un groupe social, notamment des groupes émergents. Les signes exprimés par ces groupes servent alors à baliser des espaces sociaux. Enfin, pour échapper à la multiplicité de ces approches, une solution consiste à prendre les émotions comme des éléments de discours, laissant à la personne qui éprouve le soin de dire ce qu'elle ressent. Cet étiquetage par soi peut toujours conduire à des catégorisations révélatrices de la manière dont sont pensées les émotions, l'étiquette choisie permettant ainsi de mieux comprendre comment un individu, ou un groupe social, considère ce que sont les émotions.
- 7 Dans l'approche biologique, les émotions sont des manifestations d'un trouble entre l'état du corps à un instant donné et un aléa issu de l'environnement ou de la réminiscence d'une pensée. L'émotion indique un ajustement du corps se préparant à agir pour surmonter le trouble. Par exemple, la vue d'un prédateur ou d'une proie. Cette théorie est d'inspiration darwinienne : les mécanismes émotionnels seraient le fruit d'une adaptation rendue nécessaire par les conditions de survie des primates. Ainsi l'amygdale fait-elle le tri entre des percepts ancrés dans la phylogenèse, comme l'instinct de survie à l'égard du vide ou du feu, et ceux acquis au cours de l'ontogenèse, telles les dispositions et expériences individuelles, faisant que chacun ne développe pas la même phobie du vertige. Les émotions sont donc des fonctions, associées à des zones cérébrales, qui s'activent avec plus ou moins d'intensité selon l'évaluation des situations. Le stress résulte ainsi d'une estimation que l'on ne dispose pas ou plus des ressources nécessaires ni pour fuir ni pour combattre. À cet égard, les réactions émotionnelles — dont les expressions du visage — seraient communes à toute la classe

des mammifères. Les émotions ne sont plus alors des signes qu'on articule ou interprète, mais des signaux qui s'exécutent et se mesurent.

- 8 De manière globale, la théorie des émotions basiques (*basic emotion theory* ou BET²) offre un cadre notionnel aisément transposable dans l'univers informatique. Elle gomme la complexité des interprétations culturellement situées en réduisant l'heuristique à un certain nombre de signaux discrétionnaires — la peur, la joie, la colère, la tristesse et le dégoût³. Pour chacune de ces émotions, elle identifie des marqueurs somatiques et des degrés dans l'intensité d'activation. Par exemple, plus la commissure des lèvres remonte, plus la joie est grande. Les psychologues Ekman et Friesen (1969) ont modélisé toutes les combinaisons de micro-expressions faciales en leur associant l'une des émotions basiques au sein du *Facial Action Coding System* (FACS). Par exemple, l'activation de trois sous-ensembles musculaires — le nez, les commissures, la lèvre inférieure — indique le dégoût. Ekman, le principal promoteur de la BET, reconnaît que l'expression émotionnelle est dépendante du contexte culturel, mais que cette contingence peut être dépassée par une approche biométrique. Dans *Affective computing*, l'ouvrage fondateur de l'informatique affective dans lequel Ekman est l'auteur le plus appelé (33 occurrences), Rosalind Picard le cite en ces termes : « La condition *sine qua non* pour [interpréter] une émotion ne devrait pas être un signal panculturel⁴ (Ekman, 1992) » (Picard, 2000, p. 32). À la question de savoir, une ligne après, comment un ordinateur pourrait reconnaître les émotions si tout le monde les exprime de manière différente, elle répond s'attendre à ce que des motifs similaires puissent être reconnus par une machine afin de « regrouper les individus dans des catégories⁵ » (*ibid.*, p. 33). Des programmes informatiques automatisant la reconnaissance visuelle pourraient dès lors repérer sur les visages pixélisés les déplacements de points spécifiques (les yeux, le menton, le nez, la commissure des lèvres, etc.) et par calcul algorithmique apposer une catégorisation émotionnelle. Ainsi prend naissance la mise en données des émotions, avec une double réduction en variété (cinq émotions) et en analyse par une mesure unique de l'intensité d'activation.
- 9 Si le terme *sentiment* est absent de l'ouvrage de Picard (contre 2 520 occurrences de *emotion*), l'analyse de sentiment suivra en s'appliquant cette fois à analyser des données non plus biométriques, mais sémantiques, dont le résultat sera toujours réduit à une valeur quantitative (positive ou négative).
- 10 Picard propose plusieurs cas d'usage potentiels de l'informatique affective : l'assistance aux personnes autistes, la rétroaction avec des consommateurs, la mesure d'audience, le développement personnel, la pédagogie, la recommandation de contenus, l'avatar conversationnel personnalisé, des objets connectés (*sensitive toys*). Elle conclut son chapitre ainsi : « Il y a bien sûr plein d'autres applications, c'est juste un échantillon de qu'il est possible de faire maintenant et à l'avenir⁶ » (*ibid.*, p. 111).

Donner des affects pour produire des données

- 11 La promesse faite par l'informatique affective repose sur le postulat que l'expression émotionnelle est toujours là, que le visage, la voix, le corps de manière générale, ne peuvent mentir et que ce signal peut être encodé/décodé par captation biométrique ou linguistique. Ainsi l'émotion est (une) donnée. Et c'est cette double considération épistémologique et pratique qui rend possible la *datafication* des expressions affectives en ligne.

- 12 Pour autant, le recours à la BET et à l'apprentissage machine par l'informatique affective ne permet pas d'intégrer les éléments contextuels comme l'histoire interactionnelle ou ceux situés comme les conditions d'émergence de l'expression émotionnelle. L'approche philosophique de Spinoza que nous privilégions dans nos travaux a le mérite de précéder les lectures psychologisantes des émotions, en proposant une acception plus ouverte à la relation, et donc plus compatible avec des approches en communication. Dans *Éthique* (Spinoza, 1677), il n'est pas question d'émotions, mais d'affects présentés comme l'idée d'une transformation dans le rapport entre corps et esprit (Deleuze, 1988). L'affect résulte d'une expérience de l'altérité, qui encapacite l'individu ou au contraire limite sa puissance d'agir. Cette puissance augmente si l'affection est adéquate avec la disposition individuelle. Dans cette perspective, peu importe de savoir si l'individu fait l'expérience de la joie ou de la colère, puisqu'il est affecté. C'est donc cette affection qui vise à être répertoriée par les bases de données du Web social, de manière à construire sur elle de nouvelles affections qui prendront la forme de recommandation de contenus, souvent publicitaires. Et pour intensifier l'émergence de ces affections, et donc des recettes publicitaires, leur circulation est potentialisée par la mise à disposition de fonctionnalités affectives (Alloing et Pierre, 2017). Notre hypothèse est alors que, étant donné sa visibilité dans le champ académique, seul le modèle de la théorie universelle des émotions a été mobilisé pour soutenir cette promesse affective, sans même le remettre en question, ni dans ses fondements épistémologiques ni dans ses effets psychosociaux. Chemin faisant, le recours à un modèle inadéquat de mise en données des affects — la FACS — produirait des affections inadéquates dans les recommandations de contenus. La valeur économique de ces recommandations ne tient qu'au fait que le modèle n'est pas concurrencé par une autre approche et qu'il est performant. Au cours du dernier trimestre 2024, les recettes publicitaires s'élèvent à 27 milliards de dollars américains chez Alphabet⁷ et à 15 milliards chez Facebook⁸.
- 13 La mise en données du réel est pourtant un processus largement connu et critiqué (Mejias et Couldry, 2019 ; Chignard et Benyayer, 2015). Elle repose sur « la nécessité, et pas seulement la possibilité, de traduire en données chaque variation dans la trame de l'expérience humaine afin de mieux la calculer et la traiter⁹ » (Couldry, 2019). Toutefois, la rencontre entre *datafication* et émotion — pourtant au cœur de cette trame expérientielle — a été peu abordée jusqu'à présent. À la suite des comportements relevant d'une quantification de soi (Pharabod, Nickolski et Granjon, 2013 ; Dagiral *et al.*, 2019a ; Dagiral *et al.*, 2019b), plusieurs travaux s'intéressent aux émotions que provoque la captation des émotions, notamment quand ce processus se déroule au quotidien et conditionne les routines ordinaires (Kennedy, 2018). Si, la plupart du temps, les analyses portent sur des usages psychothérapeutiques ou technopédagogiques (Belli, 2016 ; Jain *et al.*, 2024), des perspectives autres se font entendre. Ainsi l'application *Emotion Tracker* permettrait-elle d'extraire la mélancolie d'une perspective psychothérapeutique (Janasik-Honkela, 2018). Toutefois, ce sont surtout les effets de la *datafication* qui sont analysés, souvent exprimés en termes de fatigue (*data disaffection* chez Grover, Widera et Ananny, 2025).
- 14 L'étude que nous proposons ne s'arrête pas à la couche technique de la *datafication* des émotions. En montrant quelles sont les références scientifiques de cette architecture, nous proposons surtout une mise en relation avec la promesse accompagnant ce processus. Notre corpus étant constitué de travaux financés par les plus grandes

plateformes actuelles du Web social, nous proposons dès lors de montrer comment s'installe une représentation dominante où les émotions comme leur capacité à nous affecter tiendraient un rôle stratégique dans leurs modèles d'affaires au sein desquels la performance est portée par le modèle algorithmique des architectures informationnelles.

Méthode et principaux résultats

- 15 Parallèlement aux recettes publicitaires mentionnées précédemment, les dépenses mensuelles en R-D de Meta étaient de plus de 4 milliards de dollars américains, et autant chez Alphabet¹⁰. S'il est nous est impossible d'évaluer la part de ces investissements consacrée aux questions émotionnelles et affectives, il est tout de même possible d'en saisir certaines retombées au travers des publications scientifiques et des dépôts de demande de brevet. Plus que le reflet fidèle de ce que ces entreprises font réellement pour produire des données sur les affections de leurs usagers et usagères, nous émettons l'hypothèse que cette littérature scientifique et industrielle constitue avant tout des discours mis en circulation pour démontrer des capacités d'innovation en la matière. S'adressant ainsi tant aux investisseurs qu'aux concurrents, ces publications mettent au jour ce que ces industriels du numérique disent vouloir faire de nos émotions et de nos affects. À cela s'ajoute l'ensemble des textes promotionnels et des prescriptions mis à disposition des usagers et usagères de leurs plateformes. Il y est démontré l'intérêt de faciliter ou de mesurer l'expression émotionnelle des publics, particulièrement en termes de performance commerciale ou communicationnelle.
- 16 Nous avons ainsi constitué trois corpus, le premier de publications scientifiques, le second de brevets et le troisième de prescriptions. Notre choix s'est porté sur quatre entreprises propriétaires d'une ou plusieurs plateformes de médias sociaux : ByteDance (TikTok), X Corp. (X ex-Twitter), Meta (Facebook, Instagram) et Google (y compris YouTube). Afin de collecter nos corpus, nous avons systématiquement eu recours aux mêmes mots-clés propres au champ émotionnel (*sentiment, emotion, mood, feeling, joy, sadness, empathy, etc.*).

Collecte et analyse des publications scientifiques

- 17 Pour constituer le corpus de publications scientifiques, nous avons utilisé pour chacune de ces plateformes, au 1^{er} septembre 2023, leurs sites dédiés aux équipes de recherche, comme research.google ou research.facebook.com. Une première recherche nous a permis d'éliminer ByteDance (aucune publication) et X (une publication). Nous avons donc gardé Méta (30 publications) et Google (28 publications). Les articles de Meta ont été publiés de 2009 à 2023 et ceux de Google l'ont été entre 2005 et 2022. Pour nous assurer de ne pas avoir omis de publications, nous avons utilisé les mêmes mots-clés associés aux noms des auteurs (121 pour Google et 130 pour Meta) avec le logiciel Publish Or Perish (Harzing, 2007). Nous avons ensuite utilisé une grille afin d'analyser de manière systématique les publications selon 14 items comme le champ de recherche, l'orientation de l'article (théorique, applicatif, méthodologique), ses objectifs ou encore les données collectées.

- 18 Avec le logiciel VOSviewer (Van Eck et Waltman, 2011), nous avons procédé à une analyse lexico-sémantique des titres et résumés de chaque article afin de visualiser les cooccurrences des termes ou des expressions (terme1_terme2) les plus présentes dans chacun des corpus Meta et Google. Cette analyse nous permet d'une part de souligner l'association de termes du champ lexical affectif avec des intérêts des plateformes et d'autre part de fournir une vue d'ensemble des thèmes, objets de recherche et sujets discutés dans les corpus analysés. De plus, cela nous permet de développer des hypothèses quant aux possibles orientations de R-D des plateformes tout en mettant en exergue certains mécanismes propres à la *datafication* des affects (données utilisées, champs applicatifs, etc.). Les résultats de cette analyse sont principalement des éléments de démonstration et gagnent en pertinence par les approches qualitatives déployées en complément. Ici, donc, leur valeur est avant tout heuristique.
- 19 Pour Meta, nous avons gardé les 22 expressions et termes les plus redondants (c'est-à-dire présents plus de 5 fois) et les plus pertinents (suppression des adjectifs, pronoms, chiffres, etc.). VOSviewer nous a ensuite permis d'obtenir des *clusters* terminologiques, c'est-à-dire des regroupements de termes avec le plus haut taux de cooccurrence dans notre corpus. Ainsi, lorsqu'un *cluster* associe les termes A et B, plusieurs documents du corpus où ces termes sont cooccurents traitent des sujets A et B. Pour Meta (figure 1), une partie de la littérature traite des aspects sociaux des émotions (*cluster 2*), avec des termes centraux comme *friend*, *post* ou *response*, et surtout le choix du mot *feeling* (*sensation* en français) pour traiter du champ affectif. Le *cluster 1* regroupe des articles traitant spécifiquement de l'émotion et qui s'orientent vers des questions informatiques (*dataset*, *task*) ou applicatives (« image », « texte », *video*, *hate speech*).

Figure 1. Regroupements (*clusters*) des 22 termes et expressions les plus redondants dans le corpus d'articles scientifiques Meta selon VOSviewer

22 items (2 clusters):
Cluster 1 (13 items)
dataset
emotion
hate speech
human
image
meme
performance
state
task
text
touch
video
work
Cluster 2 (9 items)
comment
feeling
friend
person
post
response
social capital
social comparison
social medium

20 Pour Google, 22 termes et expressions ont aussi été retenus (figure 2) et 4 *clusters* émergent, ce que l'on peut interpréter comme une plus grande variété d'orientations de recherche dans les articles du corpus. Sans revenir sur chacun d'entre eux, nous pouvons souligner que :

- le *cluster 1* regroupe des articles qui associent l'*emotional facial expression* d'un visage (*face*) tel un sourire (*smile*) comme un moyen d'évaluer un effet en contexte ;
- ceux du *cluster 2* associent l'expérience des plateformes avec la question de la haine, des attaques ou du harcèlement (*harassment*), particulièrement lorsque cela concerne les créateurs de contenus ;
- ceux du *cluster 4* associent le *sentiment analysis* aux textes, aux applications, aux produits (comme Google Business).

Figure 2. Regroupements (*clusters*) des 22 termes et expressions les plus redondants dans le corpus d'articles scientifiques Google selon VOSviewer

22 items (7 clusters):
Cluster 1 (5 items)
context
effect
emotional facial expression
face
smile
Cluster 2 (5 items)
attack
creator
harassment
hate
platform
Cluster 3 (4 items)
child
emotional expression
narrative
storyface
Cluster 4 (3 items)
product
sentiment analysis
text
Cluster 5 (2 items)
application
problem
Cluster 6 (2 items)
developer
experience
Cluster 7 (1 item)
task

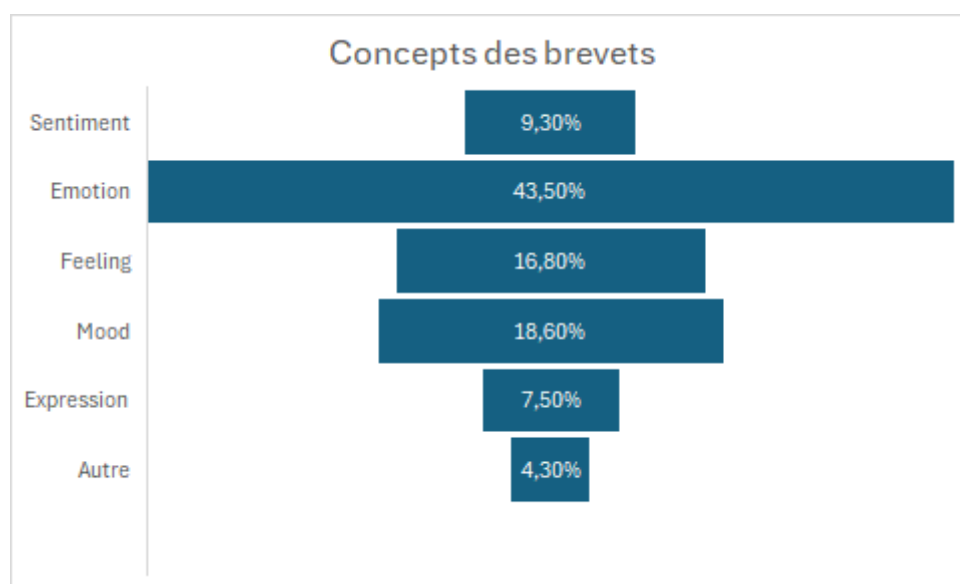
21 En somme, les deux entreprises vont tout d'abord orienter leurs recherches vers des problématiques propres aux besoins et fonctionnalités de leurs plateformes. Pour Google et son modèle documentaire construit autour de la hiérarchisation de contenus informationnels produits par les internautes ou certaines organisations, la question de l'analyse des sentiments et des opinions est ainsi centrale (Lu *et al.*, 2020 ; Sun, Liu et Prasad, 2019 ; Li *et al.*, 2014 ; Täckström et McDonald, 2011a, 2011b ; Councill, McDonald et Velikovich, 2010 ; McDonald *et al.*, 2007 ; Shafran et Mohri, 2005). Pour Meta, dont le modèle repose sur l'instrumentalisation des graphes relationnels entre usagers et usagères pour affiner la recommandation de contenus que ce soit sur Facebook ou Instagram, comprendre les émotions des publics au travers de leurs publications est une occasion de favoriser leur socialisation (Burke, Marlow et Lento, 2010 ; Kramer, 2012 ; Burke et Develin, 2016 ; Scissors, Burke et Wengrovitz, 2016 ; Burke, Cheng et De

Gant, 2020). Comme nous le verrons, ces orientations s'appuient néanmoins sur des modèles techniques communs aux deux entreprises.

Corpus et analyse des brevets

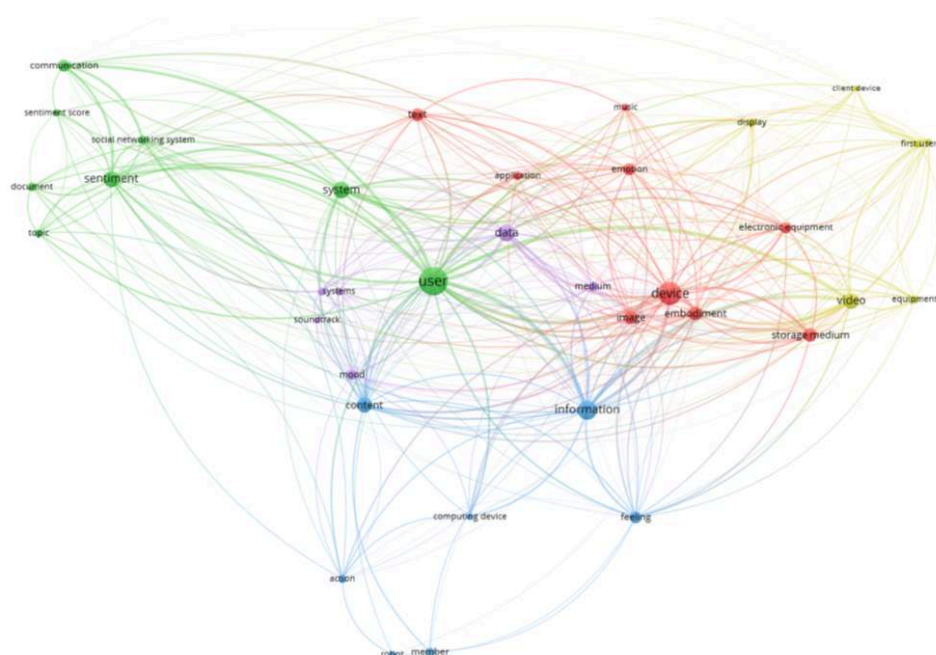
- 22 Notre corpus de brevets a été collecté avec le logiciel Patent2net (Reymond et Noyer, 2022), de janvier 2012 à septembre 2024 à partir des mêmes mots-clés que pour les articles (figure 4) : Meta (n = 20), Google¹¹ et ses filiales (n = 21), ByteDance (n = 17) et X/ Twitter (n = 5). Nous avons par la suite effectué un traitement lexico-sémantique semblable à celui des publications scientifiques. L'analyse qualitative a quant à elle été effectuée au moyen d'une grille d'analyse reprenant les mêmes items que ceux des publications scientifiques.
- 23 Dans ces brevets, différents concepts du champ affectif sont abordés (figure 3). Parmi eux, 43,5 % présentent ainsi des inventions capables de reconnaître les émotions, de les analyser ou même d'influer sur elles, là où plus de 15 % vont s'intéresser aux ressentis (*feeling*) et à l'humeur (*mood*). Un plus petit volume de brevets décrit des dispositifs centrés sur les sentiments et les expressions (faciales ou textuelles) des usagers et usagères des plateformes.

Figure 3. Principaux concepts discutés dans les brevets déposés par les quatre entreprises de notre corpus (43,5 % des brevets traitent des émotions)



- 24 Toujours à l'aide de VOSviewer, nous avons identifié cinq *clusters* terminologiques (liens de proximité entre les 36 expressions ou termes les plus fréquents) qui mettent en relief les grandes orientations de ces brevets à partir de leurs titres et de leurs résumés (figure 4).

Figure 4. Cartographie lexico-sémantique des 36 expressions et termes les plus redondants dans le corpus de brevets



- 25 On peut voir au centre de cette cartographie le terme *user*. Si mise en données des émotions il y a, alors celle-ci s'alimente des usages des publics des plateformes de ces entreprises. De même, chaque concept du champ affectif correspond à un *cluster* particulier, au sein duquel il est associé à des termes dénotant des dispositifs techniques et des finalités applicatives. Dans certains brevets, les images et la musique sont présentées comme en mesure de provoquer des émotions dont la captation devient alors possible par des équipements spécifiques (*electronic equipment*). La notion de sentiment est quant à elle associée à celle de communication, de réseaux sociaux numériques, de documents, avec pour possible finalité des inventions présentes dans ce *cluster* la production de *scores* de sentiments. Le ressenti (*feeling*) est généré par des interactions avec les contenus et les informations produits en ligne par les dispositifs informatiques et par les robots dont les brevets associés à ce *cluster* se font forts de décrire le développement. Enfin, certains brevets conçoivent l'humeur (*mood*) comme la conséquence de l'exposition quotidienne de certains publics aux médias et à la musique. Un dernier *cluster* laisse entrevoir des inventions proposant des modes personnalisés d'affichage de vidéos en fonction des affections des spectateurs et spectatrices. Ces brevets décrivent donc des inventions qui diffèrent en fonction des types d'émotions qu'elles postulent pouvoir transformer en données puis traiter. Que ce soit dans les brevets ou les publications scientifiques, la production, le traitement et la captation de données affectives reposent sur les interactions et les contenus des usagers et usagères. Pour stimuler cette production émotionnelle (commentaires, photos, vidéos, messages vocaux, clics, etc.), les mêmes entreprises développent un ensemble de prescriptions et de discours à destination non seulement de ces mêmes usagers et usagères, mais aussi des annonceurs publicitaires et créateurs et créatrices de contenus.

Corpus et analyse des prescriptions et discours promotionnels

- 26 Au sein des sites de relations de presse et de prescriptions d'usage de ces mêmes quatre entreprises¹², nous avons identifié la présence des mots-clés voulus à l'aide d'une requête (*site*) sur le moteur de recherche sur Google.com : Meta (n = 86 pages Web), Google/YouTube (n = 32), ByteDance (n = 124) et X (n = 115). Pour ce corpus, nous avons uniquement appliqué une analyse qualitative thématique à partir d'une grille de codage libre, puis un processus d'analyse inductive (Thomas, 2006). En effet, les brevets et dans une moindre mesure les publications scientifiques reposent sur une organisation des propos relativement similaires, avec une terminologie circonscrite. Les prescriptions de chaque plateforme quant à elles varient fortement tant dans leur mise en page que pour le choix de termes, parfois inventés pour distinguer une plateforme de l'autre¹³. De ce processus ont émergé quatre grandes catégories thématiques permettant de mettre en relief les représentations des affects que véhiculent ces plateformes, ainsi que ce qui est attendu de leur instrumentalisation. Tout d'abord, activer les émotions des publics permet de *créer un lien* avec eux, puis de *générer du profit* et des données utiles à la recherche (marketing et scientifique). Pour ce faire, il convient d'utiliser les *fonctionnalités affectives* proposées par ces plateformes (ajout de filtres, de sons, d'images, d'émoticônes, etc.), de valoriser les identités des publics (nationales, de genre, sexuelles, etc.) ou de communiquer autour d'événements réunificateurs et commerciaux (Saint-Valentin, Noël, etc.). Afin d'éviter que toutes ces affections n'entraînent des conséquences délétères, et pour prendre soin des usagers et usagères tout en assurant leur sécurité, les plateformes soulignent l'importance d'*encadrer les expressions émotionnelles*, particulièrement les contenus haineux. Tout cela permet alors d'avoir des plateformes qui favorisent la *diffusion de la joie* partout dans le monde ! À la lecture de ces corpus, et par leur analyse, se dessine donc un ensemble de techniques et de méthodes de mise en données de nos affections et ressentis. Les ambitions des plateformes semblent à la (dé)mesure des discours sur les émotions qu'elles charrient au travers de ces diverses publications.

Datafication des émotions et instrumentalisations des affections : ce que les plateformes disent vouloir faire de nos émotions

- 27 Nos différents corpus nous offrent une vision partielle, mais néanmoins précise, de ce que les plateformes que nous étudions revendiquent et ambitionnent de faire avec les affects de leurs usagers et usagères. Plutôt que d'en proposer une description puis une longue analyse systématique, nous choisissons ici de donner une vue d'ensemble des processus à l'œuvre dans ces stratégies de mise en données et de traitement de ce que les plateformes qualifient d'émotion.
- 28 Un premier constat est que ces entreprises ne s'embarrassent pas d'une définition précise des objets qu'elles prétendent pouvoir capter, mettre en données et traiter. La quasi-totalité des brevets ne définit pas ce qu'est une émotion, se contentant d'en citer la gamme (peur, joie, tristesse, etc.) ou de fournir une explication floue : « L'émotion peut être un état mental associé à une activité physiologique et peut être influencée par des conditions internes ou externes¹⁴ » (US20190197073A1). Si une approche aussi peu

étayée peut favoriser le dépôt d'inventions au spectre applicatif très large, il est plus étonnant de ne pas trouver de discussion aboutie dans les publications scientifiques de ces entreprises. Par exemple, seuls dix articles sur trente publiés par les équipes de Google définissent les émotions (Kelley *et al.*, 2021 ; Egelman *et al.*, 2020 ; Lu *et al.*, *op. cit.* ; Sun, Liu et Prasad, *op. cit.* ; Liu, Sosik et Singh, 2018 ; Ryokai, Raffle et Kowalski, 2012 ; Titov et McDonald, 2008 ; McDonald *et al.*, *op. cit.* ; Cui, Mittal et Datar, 2006 ; Shafran et Mohri, *op. cit.*). Comme pour Meta (Kramer, 2012 ; Burke et Develin, *op. cit.* ; Wang, Burke et Kraut, 2016 ; Shukla, Petridis et Pantic, 2021 ; Rognon, Stephens-Fripp *et al.*, 2022 ; Jegorova, Petridis et Pantic, 2023), les émotions y sont principalement discutées par le prisme des travaux de Paul Ekman et sont uniquement caractérisées par leur valence positive, négative ou neutre et leur intensité, ou *arousal* (Kreuk *et al.*, 2021). Quant aux prescriptions des plateformes, elles tiennent pour acquis qu'une émotion est avant tout un effet produit sur les individus, dit autrement, une affection.

- 29 En première lecture, les émotions apparaissent donc comme des données au sens le plus réducteur : elles sont universelles, leur description est univoque et elles ne prêtent pas à discussion ou à débat. Ne reste plus alors qu'à les structurer pour produire de la valeur. Nous allons dès lors discuter dans la partie suivante ce que pourrait être cette valeur, puis les moyens mis en œuvre pour la produire et les conséquences probables de cette production, avant de mettre en exergue des questionnements éthiques nécessaires.

Promesse et représentation des « données affectives »

- 30 Suivre l'innovation depuis des articles et brevets jusqu'aux prescriptions faites aux marques, aux créateurs et créatrices de contenus ainsi qu'aux usagers et usagères nous permet d'élaborer la chaîne de valeur des technologies affectives telles que les imaginent les plateformes.
- 31 Les émotions exprimées sur lesdites plateformes sont évidemment la donnée centrale. Parmi les brevets collectés, 51 % proposent de les identifier afin d'anticiper les réactions possibles et d'activer des recommandations personnalisées. Ces recommandations concernent les choix d'emoji ou de *sticker* à employer selon le texte en cours d'écriture, et plus largement les informations et autres contenus. Les modes de classification envisagés sont encore hétérogènes (par intensité, par polarisation binaire, par agrégation), mais les sources couvrent autant les publications en ligne que les messageries privées. Les brevets, dans 17 % des cas, visent l'analyse de contenus multimodaux (image, musique, vidéo) à des fins de sous-titrage, de recommandation de musique pour accompagner un texte ou d'élaboration de collections (*playlist*). Les émotions sont présentées comme un flux d'information reflétant des changements d'humeur chez l'utilisateur et l'utilisatrice, et 41 % des brevets visent à adapter l'interface ou le parcours-utilisateur selon ces variations. Enfin, un brevet sur quatre encourage l'expression émotionnelle individuelle dans la messagerie, lors de diffusion en direct (*streaming*), dans des événements virtuels, dans les fils personnalisés (*newsfeed*), par des indicateurs sémiotiques ou d'autres fonctionnalités.
- 32 Les cas d'usage sont aussi variables, particulièrement dans les articles scientifiques. Les plateformes se positionnent sur des filières technologiques complémentaires, comme les technologies sensorielles chez Facebook et la robotique chez Google. Dans ce dernier cas, il s'agit de développer des robots thérapeutiques pour accompagner les personnes

ayant un trouble du spectre autistique. De manière générale cependant, les finalités sont d'abord centrées sur l'expérience proposée aux utilisateurs des médias sociaux. Sur cette base individuelle, les plateformes envisagent aussi la production de données agrégées : les avis clients chez Google ou des statistiques sur le « Bonheur National Brut » chez Facebook. Twitter est la seule plateforme à inciter explicitement les marques à provoquer des émotions pour obtenir des données afin de mieux analyser les audiences (A2)¹⁵.

- 33 Les créateurs de contenus restent la figure principale dans les préoccupations que les plateformes expriment au travers de leurs prescriptions d'usage. Il peut s'agir aussi bien de créateurs indépendants que d'employés dans des services ou agences de communication marketing. Les discours qui leur sont destinés dans des pages dédiées listent d'abord toutes les fonctionnalités affectives qu'ils peuvent mobiliser selon leurs objectifs. En tant qu'énonciateur, les émojis servent « à indiquer les émotions par rapport au contenu d'une publication » (A3), rappelle Facebook, tandis qu'ils offrent « la possibilité à [leur] audience de partager son ressenti » (A4) sur Instagram. Sur cette plateforme, les filtres visuels sont aussi mis de l'avant dans les prescriptions, afin d'« associer une humeur au contenu » (A5). Sur Twitter, les filtres apportent du « style et de l'humeur » (A6) aux photos. Le *storytelling* est aussi encouragé, et Meta prône le développement d'histoires (*stories*) de marque initiées par les émotions (*emotion-led brand stories*, A7). Facebook est encore plus explicite : il s'agit de faire du « *storytelling* avec son cœur » (A7).
- 34 Dans cette veine, les créatrices et créateurs de contenus sont invités à voir les événements comme autant d'occasions émotionnelles de générer du profit. Au moment où nous écrivons cet article, c'est la saison de la Saint-Valentin (A8), et les conseils sont nombreux sur les pages consultées. La rentrée des classes ou les Jeux olympiques sont aussi des événements propices. L'invitation est claire chez Facebook : « Découvrez comment les équipes de marketing de performance peuvent *parler le langage de l'amour* et développer du contenu publicitaire qui suscite l'émotion et incite à l'action » (A9) (nous soulignons). « *Love always* », nous dit TikTok à cette occasion (A10). Le Mois des fiertés apporte aussi des possibilités. Elles sont essentiellement commerciales chez Meta, qui donne des conseils pour « tirer parti des fonctionnalités d'achat dans votre propre boutique pendant cette saison de la fierté » (A11). L'approche est plus communautaire chez TikTok, qui invite à le rejoindre lors de la « *#ForYourPride TikTok celebration* » (A12) ; la plateforme considère ainsi que les membres de la communauté LGBTQIA+ ou de *#BlackTikTok* sont les plus créatifs et les plus positifs (A13). Twitter/X tient un discours tout autant inclusif (au moment de notre collecte). TikTok se distingue dans ces discours en accordant — et en préconisant de faire — beaucoup de place aux témoignages de ses créateurs et créatrices : l'une d'entre elles déclare ainsi se sentir « plus confortable et moins seule » (A14) parce que TikTok lui permet de revendiquer son identité culturelle. Dans les pages consultées, depuis le Québec, les « *French-speaking Canadians* » (A15) sont aussi valorisés.
- 35 Le sentiment d'appartenance à une communauté est essentiel chez Meta, dont la mission est, depuis 2017, de « donner aux gens le pouvoir de construire des communautés et de rassembler le monde¹⁶ ». TikTok l'énonce de manière similaire : « Rassembler est le maître-mot » (A16). Pour y arriver, les auteurs des pages-conseils proposent un exercice simple de texte à trous (A9), dans une page intitulée « Du Cœur au Panier » d'achat (figure 5).

Figure 5. Impression d'écran de la page Facebook Business Insights

2. Dites-le avec le cœur

Une fois que le bon territoire émotionnel a été identifié, il est temps de le transformer en l'expression d'un bénéfice. L'objectif ici est de le dire avec le cœur, de développer un langage d'amour entre la clientèle et le produit. En d'autres termes, communiquez sur la manière dont le produit peut améliorer la vie d'une personne. Voyez cette étape comme un exercice à trous où le produit, l'émotion et le bénéfice émotionnel sont clairement exposés.

_____, car _____
Produit émotion bénéfice émotionnel

- 36 Les données émotionnelles sont aussi valorisées par des traitements algorithmiques peu explicités par les plateformes. Quand ces traitements sont mentionnés par TikTok, c'est à l'intérieur des solutions créatives que la plateforme dit offrir à l'industrie¹⁷. Finalement, ce sont surtout les actions de surface qui sont préconisées : les conseils dispensés insistent sur les effets des images (« Jouez sur la corde sensible avec des photos de produits captivantes », A8, Instagram), des vidéos (« connexion émotionnelle », A17, Twitter), du son (« *a game changer* », A18, TikTok).
- 37 C'est donc bien la performance affective des usagers et usagères comme des annonceurs qui est visée ici, contrairement aux brevets et à la littérature scientifique qui décrivent les mécanismes permettant de stimuler ou de capter de manière massive et automatisée cette performance. Il nous apparaît que l'application de ByteDance est celle qui s'inscrit le plus dans cette démarche — discursivement en tout cas. En effet, TikTok se présente comme un « *total mood booster* » (A19) dont le seul objectif est la joie : « apporter la joie [...] déclencher la joie [...] trouver la joie » (A20, A21, A22). De même que Reddit se présentait comme la page d'accueil du Web, TikTok se voit comme le dernier coin de soleil d'Internet¹⁸. Nous interprétons cette posture comme une manifestation de l'idéologie de « l'énergie positive », une relecture contemporaine du confucianisme par le parti communiste chinois et qui se retrouve également dans d'autres applications comme WeChat (Rocher, 2024).
- 38 L'objectif est le même, mais l'espace de réalisation est plus restreint chez Meta : « Grâce aux boutiques, nous voulons offrir aux gens un espace sur lequel découvrir les joies du shopping, loin des corvées de courses » (A23), peut-on lire sur Instagram. L'empathie est ainsi pensée comme un « sentiment essentiel [pour] attirer une nouvelle clientèle et la fidéliser » (A24). La surcouche émotionnelle viendrait conforter des effets de levier stratégique pour les marques. Les visions behavioristes de l'émotion resurgissent ici. Pour Meta, « les achats sont guidés par l'émotion et confirmés par la raison » (A25), sur Instagram il faut transformer l'intérêt en désir puis en vente¹⁹(A26). « Les émotions dirigent les ventes » (A27), selon Twitter, et « une découverte joyeuse mène à des achats inspirés » (A28).
- 39 Ces intentions de positivité, d'expérience optimisée, de conversion d'achat, s'accompagnent d'une condition *sine qua non* spinoziste. Pour maintenir la joie, il faut réduire la haine. Ces efforts sont plus manifestes dans les brevets et les articles, où la reconnaissance des sentiments négatifs est largement explicitée. En surface, c'est d'abord la bienveillance qui est présentée afin de mieux accueillir et protéger non seulement les créateurs et créatrices de contenus, mais aussi les salariés²⁰. À leur sujet, TikTok parle d'encapacitation (*empowerment*, A30), Twitter de santé mentale (A31), YouTube — pendant la pandémie — de soin de soi (*self care*, A32). Pour consolider

l'espace sain qu'affichent les plateformes (*safe space*, A33), des fonctionnalités de dénonciation et de protection anticipée (*safety mode*, A34) sont mises à disposition des usagers et usagères (*Address anti-Black sentiments when you see them*, A35, Twitter), des campagnes anti-haine sont déployées (*Swipe Out Hate*, A36, TikTok) et les comptes promouvant des discours de haine sont « démonétisés ».

- 40 S'agissant de promesses et de représentations, la stratégie affective des plateformes relève pour l'instant d'un discours destiné aux créateurs et créatrices de contenus pour les convaincre de la pertinence de stimuler et de faire circuler des affects, les leurs et ceux de leurs publics. Or c'est bien parce que la promesse — par la simplicité de sa formule — minore le processus de traitement des expressions émotionnelles et des affections que la *datafication* peut plus facilement être invisibilisée. Il nous apparaît alors nécessaire d'étudier plus en profondeur les traitements inhérents à cette mise en données des émotions.

Production et traitement des « données affectives »

- 41 Dans nos corpus de publications scientifiques et de brevets, les moyens mis en œuvre dépendent du spectre affectif qui est mobilisé pour produire des données dites affectives à des fins de développement de dispositifs spécifiques. Deux grandes catégories de données sont utilisées : celles issues des productions des usagers et usagères de plateformes numériques et, dans une moindre mesure, celles biométriques en provenance de dispositifs portatifs (téléphones, casques de réalité virtuelle).
- 42 Dans la première catégorie, les éléments collectés proviennent des plateformes dont les entreprises étudiées sont propriétaires. Les équipes de recherche vont ainsi constituer des corpus à partir de publications sur Facebook (Kramer, 2010 ; Burke, Marlow et Lento, *op. cit.* ; Kramer, 2012 ; Burke et Develin, *op. cit.*), d'avis et commentaires sur *maps.google.com* (Blair-Goldensohn *et al.*, 2008) ou *shopping.google.com* (Cui, Mittal et Datar, *op. cit.*). Certains brevets décrivent plus précisément les données utilisées pour l'identification de sentiments, comme les émojis pour Twitter (US20200073485A1), les émoticônes pour Google (US20180059885A1) ou encore le croisement entre des éléments textuels et des « informations biographiques, démographiques et d'autres types d'informations descriptives, telles que l'expérience professionnelle, les antécédents scolaires, les loisirs ou les préférences, le lieu de résidence » pour Facebook (US20180165582A1). Mais ces entreprises ne se limitent pas à leurs plateformes. Des publications Reddit (Demszky *et al.*, 2020 ; Dwivedi-Yu et Halevy, 2022), des critiques de films sur IMDB (Mesnil *et al.*, 2015) ou de commerces sur TripAdvisor (Titov et McDonald, *op. cit.*) sont aussi collectées. Il en va de même avec la reconnaissance des émotions dans les images, les vidéos ou encore le son. Les données sont produites par les applications mobiles de ces plateformes. Par exemple, pour évaluer la réception affective d'une vidéo, TikTok propose dans un brevet de collecter « des informations vocales de l'utilisateur » et « une image du visage d'un utilisateur » (CN111443794A).
- 43 Les données affectives apparaissent alors comme des données secondaires. Elles sont produites par l'agencement d'un grand volume de données primaires souvent hétérogènes, mais dont la qualification est nécessaire pour assurer par la suite une standardisation des processus de collecte, de production et d'analyse. Si, de 2010 à 2015, les travaux de recherche que nous avons étudiés mettent l'accent sur la capacité à

produire une grande diversité de données sur les affects de leurs usagers et usagères, ceux qui suivent se concentrent particulièrement sur la nécessité de structurer ces jeux de données pour favoriser la reproductibilité et la standardisation des processus.

- 44 Les entreprises vont utiliser pour leurs recherches des bases de données ouvertes proposant de grands volumes de données préqualifiées, c'est-à-dire annotées pour, par exemple, indiquer une intensité ou directement une émotion dans une image, un son, un texte. Spécifiquement constituées pour entraîner des algorithmes d'apprentissage, elles favorisent l'identification de contenus haineux (HateXplain, OffensEval), la reconnaissance des émotions sur les visages dans les images animées ou non (YouTube-8M, ImageNet) ou encore dans la voix (Interactive Emotional Dyadic Motion Capture). Le recours à ces bases de données reflète l'universalité supposée des émotions : puisqu'elles sont exprimées de manière similaire par toutes et tous, alors il est possible d'en proposer une qualification monosémique. Les bases de données sur les discours de haine sont élaborées à partir d'annotations produites avec la plateforme Amazon Mechanical Turk, les catégorisations sont basées sur la description des émotions issue des travaux de Picard et les données biométriques pour l'entraînement proviennent d'enregistrements avec des comédiens (*motion capture*). L'apprentissage machine se révèle alors utile pour industrialiser la production, puis le traitement de ces données secondaires qui sont qualifiées de « sentiments », « ressentis » ou « émotions ».
- 45 Dans les recherches, les approches par *self-supervised machine learning* sont favorisées (Shukla, Petridis et Pantic, *op. cit.* ; Jegorova, Petridis et Pantic, *op. cit.*) avec des modèles répandus comme HuBERT (Kreuk *et al.*, *op. cit.*), de même que les *convolutional neural networks* (Sun, Liu et Prasad, *op. cit.* ; Hickson *et al.*, 2019 ; Attia *et al.*, 2018 ; Marinoiu *et al.*, 2018), notamment pour l'analyse du langage naturel (Demszky *et al.*, *op. cit.* ; Lu *et al.*, *op. cit.* ; Sun, Liu et Prasad, *op. cit.*). Dans une optique de standardisation et d'innovation ouverte (Chesbrough, Vanhaverbeke et West, 2006), ces entreprises mettent à disposition leurs bibliothèques logicielles (comme PyTorch pour Meta), des bases de données affectives annotées (Dwivedi-Yu et Halevy, *op. cit.*) ou leurs propres technologies d'apprentissage non supervisé (comme Transformer chez Google). Ces modèles informatiques favorisent par la suite la production de données affectives dans des contextes spécifiques que nous pouvons identifier au travers des brevets : par Google pour définir « l'humeur d'une musique » (US8977374B1), par TikTok pour obtenir une classification émotionnelle d'une vidéo (CN111767431A) ou encore par Twitter pour labéliser émotionnellement un *tweet* (WO2022256584A1).
- 46 La *datafication* de nos affects mis en circulation sur les plateformes de médias sociaux repose donc avant tout sur des modèles qui s'autoalimentent, où les données qualifiées par l'un vont permettre de produire celles utiles à l'autre. Et c'est lorsque nous cliquons, nous écrivons, nous parlons, nous partageons nos photos que nous activons ces modèles qui vont alors déterminer ce que nous allons voir, avec qui nous allons interagir, ce qui doit et va nous affecter ou non. Zizi Papacharissi (2015) envisage les médias socionumériques comme des structures de sentiments où s'agrègent des publics affectifs par le biais de certaines fonctionnalités comme les *hashtags*. Les processus de mise en données des interactions et productions de ces publics affectifs nous permettent de dire qu'au-delà de ces structures de sentiments nous faisons bien face à des infrastructures affectives. La régulation même de leur fonctionnement repose en partie sur leur capacité à modéliser ce qui nous a affecté, va nous affecter ou pourrait le

faire. Si un message nous est présenté sur Twitter ou Facebook, une image sur Instagram ou une vidéo sur TikTok, c'est que les termes qui y sont employés, les éléments représentés, la musique, font écho à des données primaires (textes, images, etc.) ou secondaires (sentiments, humeurs, etc.) déjà traitées. Ce qui nous affecte est donc le résultat de modélisations dont les tenants et les aboutissants sont invisibilisés. Cette invisibilisation est propre à toute forme de maintenance des infrastructures informatiques (Star et Strauss, 1999). La mise en données de nos affects repose donc sur le travail des usagers et usagères et sur l'interprétation de travailleurs des plateformes qui font usage de signes affectifs comme les émojis ou les *likes*, qui expriment des sentiments dans leurs commentaires ou encore qui sourient face à une caméra. C'est par l'agrégation de ces expressions intentionnelles ou non, en réponse à une injonction ou pas, qu'il devient par la suite possible de construire des modèles informatiques aptes à simuler des émotions et à stimuler des affections. Et, dans un mouvement circulaire, ces modèles sont ensuite utilisés pour construire de nouveaux dispositifs de propagation des affections et de mise en données des affects. En somme, la *datafication* des affects apparaît comme un moyen d'évaluer et d'orienter les affections des usagers et usagères. Ces affections étant un vecteur de notre puissance d'agir, les discours produits visent à souligner l'importance de leur activation continue.

- 47 Ce recours aux données affectives ne se fait pas sans produire des effets de bord. Cela est donc particulièrement le cas lorsqu'elles sont utilisées non seulement pour stimuler des affections et orienter l'attention ou l'agir des usagers et usagères, mais aussi pour alimenter des « intelligences artificielles » qui vont par la suite automatiser la production de contenus et réguler les interactions afférentes (Stark et Hoey, 2021). La création de profils affectifs construits pour favoriser le ciblage publicitaire laisse entrevoir une question de *privacy* (McStay, *op. cit.*) : jusqu'à quel point ce qui relève de l'intime, voire de la corporalité, est utilisé dans un cadre commercial ou non ? Ensuite, la qualification même de ces données en sentiments, émotions ou ressentis laisse envisager une potentielle normalisation des expressions affectives, particulièrement lorsque celles-ci participent à la mise en visibilité algorithmique des contenus produits sur les plateformes. Les plateformes nous invitent à produire des données qui serviront à qualifier nos propres agirs, mais aussi ceux des autres. Présentées comme capables de décrire des émotions considérées comme universelles, ces données reçoivent, de la part de ces entreprises, un traitement qui réduit potentiellement la médiatisation des diversités de genre, de culture, etc., dont les expressions émotionnelles ne feraient pas écho aux standards techniques et aux présupposés épistémiques que nous avons identifiés dans nos corpus. C'est d'ailleurs pour cela qu'elles prescrivent les meilleures pratiques pour favoriser ces expressions minoritaires. Ce qui nous émeut doit s'inscrire dans les standards techniques des plateformes entraînant par là même une normalisation de nos affections, et une forme de confiscation de nos ressentis.

Ce que la *datafication* fait à nos affects : de la puissance d'agir à l'aliénation émotionnelle

- 48 Dans notre recherche, nous abordons les stratégies pour opérer la *datafication* de nos affects comme des agencements complexes, aux multiples composantes techniques et technologiques, mais aussi discursives et idéologiques. Ces infrastructures affectives

entraînent ultimement un processus d'aliénation des activités humaines, et plus spécifiquement du vécu émotionnel des individus, sous la forme de *data*. De fait, l'extraction des expériences intimes et affectives des individus en données à des fins de valorisation marchande et d'amélioration des services proposés induit une profonde distance quant à la richesse existentielle relative à ces expériences, c'est-à-dire une abstraction de leur dimension qualitative. Le terme d'*agencement* n'est pas anodin et renvoie au lexique de Gilles Deleuze et Félix Guattari, dont certains concepts apparaissent toujours d'actualité pour aborder les mutations et dynamiques du numérique.

- 49 Nous proposons plus spécifiquement de nous appuyer sur le concept d'agencement. Pour Deleuze et Guattari, un agencement présente deux dimensions ou plutôt deux facettes qui nous semblent particulièrement adaptées pour penser le mouvement de *datafication* des affects. Deleuze et Guattari qualifient d'agencement machinique le mélange complexe « de corps, d'actions et de passions, mélange de corps réagissant les uns sur les autres » et d'agencement collectif d'énonciation ce qui se rapporte à une série « d'actes, et d'énoncés, transformations incorporelles s'attribuant au corps » (Deleuze et Guattari, 1980, p. 630). Pour simplifier, un agencement collectif d'énonciation ne renvoie pas seulement à un énoncé individuel, mais bien à différents systèmes de signes (langagiers, visuels, etc.) qui peuvent se traduire en discours, images, fictions, représentations de toute sorte. C'est la raison pour laquelle nous portons, dans notre recherche, notre regard sur les discours issus des brevets, des publications scientifiques et marketing qui visent comme on l'a vu à stimuler la production émotionnelle des utilisateurs, laquelle laisse des traces textuelles, sonores ou encore visuelles. Ces discours participent à la constitution d'un imaginaire scientifique, de normes, de prescriptions et d'usages, où l'émotion apparaît précisément comme quelque chose de fondamentalement calculable et profitable. Et les traces sont traitées et captées par des infrastructures affectives dont les configurations et le fonctionnement dérivent de cet agencement collectif d'énonciation que nous cherchons à analyser. Nous en venons à présent à l'autre face de l'agencement, soit la face machinique.
- 50 À la différence de l'agencement collectif d'énonciation qui présente plus spécifiquement une dimension discursive et symbolique, l'agencement machinique porte à considérer les relations multiples et transversales entre des éléments matériels, techniques, humains, énergétiques en tant qu'ils agissent et réagissent les uns avec les autres. Le choix du terme de *machine* n'est pas anodin. Pour Deleuze et Guattari, la machine donne l'image d'un appareil mécanique constitué par des éléments hétérogènes, animés par des mouvements, des automatismes, des circulations énergétiques. Dans une telle configuration, l'humain est relégué à une position subalterne, pièce et rouage d'un agencement machinique plus large qui génère des automatismes et des répétitions ; un agencement machinique sur lequel il n'a par ailleurs aucune prise. Reprenant à plusieurs reprises les mots-clés de la cybernétique (feedback, contrôle, information), Deleuze et Guattari qualifient d'ailleurs cette relation humain-machine d'asservissement machinique pour souligner l'enrôlement et la mise au travail des fonctions perceptives, des valeurs affectives et des schèmes de pensée des individus pour assurer le fonctionnement d'ensemble. Dans notre recherche, l'agencement machinique concerne les liens et interactions entre les plateformes, les logiciels, les ingénieurs, les utilisateurs et les données que nous avons précisément

voulu cartographier. À cet égard, les brevets sont riches d'enseignement. Leur analyse nous a permis d'observer le rôle majeur des technologies algorithmiques et des modèles d'apprentissage profond dans le processus de *datafiction* du vécu émotionnel des individus. Nous proposons d'analyser cette capture de « l'énergie affective » en données (sons, images, textes) jusqu'à leur traitement grâce à une gamme de techniques d'analyse d'émotions qui sert les intérêts de ces grandes firmes technologiques que sont Meta, Google, X ou ByteDance, comme un large processus d'aliénation émotionnelle.

- 51 Prenons l'exemple d'un usager de Facebook. Son énergie affective et son vécu émotionnel vont éventuellement transparaître sur son visage ou dans ses commentaires. Des techniques d'apprentissage automatique vont détecter de la joie ou de la colère, de la peur ou de la surprise. Ces émotions circulent et s'affranchissent des situations singulières, des contextes culturels, des relations intersubjectives au cours desquels elles ont pu s'exprimer. Elles sont capturées par des dispositifs que nous avons identifiés dans les brevets et les publications scientifiques : entraînement de l'intelligence artificielle pour l'amélioration d'agents conversationnels, surveillance pour identifier les comportements toxiques et dangereux, pour prédire et suggérer des articles et services par de la publicité ciblée. Cette capture concerne l'agencement machinique d'abord, notamment lorsque cette aliénation émotionnelle sert à améliorer la performance financière de ces firmes, des technologies algorithmiques qu'elles développent, des savoirs des scientifiques qu'elles emploient, de l'activité des usagers et usagères qu'elles exploitent et qu'elles asservissent. Et l'agencement collectif d'énonciation ensuite, lorsque l'analyse et le traitement automatisés des émotions viennent nourrir une vision du social qui légitime et banalise cette surveillance intime des expressions faciales des usagers et usagères, de leurs commentaires et autres *likes* venant nourrir un « territoire émotionnel », pour reprendre l'expression de Meta.
- 52 Au sein de ces discours, les régulations juridiques et les normes éthiques pèsent peu. Dans les brevets et articles scientifiques analysés, nous n'avons observé que très peu, voire aucune considération éthique approfondie sur la production et le traitement des données affectives (Alloing *et al.*, 2025). Tout se passe comme si cette surveillance de l'intime systématiquement associée à nos usages et communications en ligne ne posait aucun problème dans la mesure où elle se présente comme une condition nécessaire à l'amélioration du fonctionnement des plateformes et au contrôle des comportements, positifs ou négatifs, c'est selon. Lorsque le terrain de l'éthique est investi, c'est généralement sous l'angle managérial de la vie privée qui amène son lot de mots-clés passe-partout, série de signifiants flottants constitutifs des agencements collectifs d'énonciation analysés : transparence, gestion responsable, principe d'explicabilité, mécanisme de gouvernance, etc. Or, suivant une perspective spinoziste, l'éthique commande de réfléchir aux conditions de possibilité d'une vie bonne et aux causes qui en bloquent l'accès par une limitation de la puissance d'agir. Dans la terminologie de Spinoza, il y a une force désirante, un pur élan, mais qui n'a cependant aucune direction donnée. C'est un désir sans objet. Dès lors, les objets que l'on essaie de poursuivre ne peuvent que venir du dehors, d'une rencontre avec des choses, des corps, des représentations. La force désirante institue des choses comme objet de désir, investissements qui sont entièrement déterminés par le jeu des affects. Comme l'a bien montré Lordon (2010), chaque société se donne des objets de désir licites. Plus précisément, les structures sociales, déterminées par des rapports de production capitaliste, configurent les désirs et prédéterminent les stratégies pour les atteindre.

Dans notre recherche, on observe que les usagers et usagères de ces plateformes entretiennent un rapport d' enrôlement avec ces firmes technologiques qui ont besoin de leur puissance d' agir affective et émotionnelle. L' enrôlement fonctionne par le désir de rejoindre ces infrastructures affectives qui multiplient les occasions de joie par l' accès aux marchandises émotionnelles comme promesse d' expériences affectives fortes et mémorables. Ces expériences constituent des moments joyeux du rapport affectif des usagers et usagères aux plateformes. Or, comme on l' a vu, ces expériences émotionnelles circulent dans un agencement machinique complexe et opaque qui tire profit de la puissance d' agir individuelle et collective. Plutôt que de servitude volontaire, il faudrait plus spécifiquement parler de servitude passionnelle pour souligner ici l' idée que les usagers et usagères de ces plateformes s' activent joyeusement sur ces plateformes pour servir un désir qui n' est pas le leur. Autrement dit, tout le problème de l' aliénation émotionnelle que nous avons cherché à étudier se présente dans les discours scientifiques, marketing et technologiques comme une forme d' aliénation joyeuse et stimulante.

Conclusion

- 53 Vous êtes témoin d' une scène qui vous semble injuste et que vous voulez partager au plus grand nombre : quelle plateforme pourrait accueillir votre indignation ? Sachant que des algorithmes peuvent décomposer ce sentiment en colère et dégoût, dans quelle mesure TikTok, plateforme de la joie, va-t-elle prioriser la circulation de votre témoignage ? Quel accompagnement musical viendra proposer Instagram pour maximiser les interactions ? Dans quelle boucle de recommandation YouTube enfermera-t-il votre vidéo ? Qualifiant votre discours de haineux, quel autre contenu proposera Facebook à vos abonnés pour modérer votre indignation ? Comment seront qualifiés à leur tour vos commentateurs tout aussi indignés que vous, plus virulents peut-être, ou désabusés, aigris, cyniques ? Finalement, quelle sera la légitimité sociale de votre indignation au regard de la qualification algorithmique des émotions exprimées dans votre message ? Si vous voulez que votre message porte loin, lisez les manuels des plateformes du Web social et apprenez à sourire, quelle que soit l' intensité de votre colère ou de votre dégoût. Au sens de Spinoza, même la Joie devient Tristesse, dans la mesure où les plateformes la rendent inadéquate.
- 54 Dans notre article, nous montrons en effet qu' au travers de leurs publications, les quatre plus grandes entreprises détentrices de plateformes dites de médias sociaux font des innovations affectives un levier de leur croissance économique. Plus que ça même, les émotions exprimées permettent la régulation des usages et de la circulation des contenus. Pour ce faire, elles déploient des stratégies de R-D visant à mettre en données nos affects pour favoriser nos affections. Cette *datafication* suppose la collecte de données primaires principalement produites par les usagers et usagères, par leurs écrits, leurs interactions, mais aussi leurs corps, puis leur traitement par des algorithmes d' apprentissage pour produire des données secondaires qui seront qualifiées de sentiments, d' émotions, d' humeurs ou encore de ressentis. Ces données affectives sont à la fois des données secondaires, puisqu' elles n' existent que par l' agencement d' autres données, et des métadonnées, puisqu' elles sont nécessaires pour affiner la description de certains objets (au sens informatique du terme).

- 55 Le choix assumé d'une appréhension universalisante des émotions, c'est-à-dire qui se défait de tout contexte, ne répond pas seulement aux impératifs techniques de leur mise en données. Il permet aux plateformes d'invisibiliser leurs modes de gouvernance. Qualifier tel propos de haineux ou tel autre de positif a une incidence directe sur sa médiatisation, sa circulation, sa discussion, et bien entendu sur la manière dont il va affecter certains individus. Il y a par conséquent lieu de s'interroger sur les implications des nouvelles politiques de modération par Twitter/X depuis son rachat ou par Meta depuis les élections présidentielles aux États-Unis en 2024²¹. La définition même de ce qui est négatif y évolue pour laisser place à des expressions autrefois qualifiées de haineuses, et les prises de parole de certains usagers et usagères autrefois fêtés pour leurs diversités sont qualifiées négativement. Ce faisant, ce qui nous affecte ne nous permet plus de persévérer dans notre être, mais uniquement de renforcer la puissance d'agir des plateformes. La mise en données de nos affects va de pair avec l'accroissement du pouvoir de celles et ceux qui l'opèrent à l'échelle industrielle.

BIBLIOGRAPHIE

- ALLOING Camille et Julien PIERRE (2017), *Le Web affectif. Une économie numérique des émotions*, Créteil, Institut national de l'audiovisuel (INA).
- ALLOING Camille et Julien PIERRE (2020), « Le tournant affectif des recherches en communication numérique. Présentation », *Communiquer. Revue de communication sociale et publique*, 28, p. 1-17.
- ALLOING Camille et Julien PIERRE (2021), « Nudges ou affordances ? Tisser la toile d'une affection distribuée », *Actes sémiotiques*, 124, DOI : <https://doi.org/10.25965/as.6751>.
- ALLOING Camille, Elsa FORTANT, Julien PIERRE, Fabien RICHERT et Rémi PALISSER (2025), « Knowing your users by heart: A critical examination of the scientific research on emotions conducted by social media platforms », *Social Media + Society*, 11(3), DOI : <https://doi.org/10.1177/20563051251355>.
- ARVIDSSON Adam (2011), « General sentiment: How value and affect converge in the information economy », *The Sociological Review*, 59, p. 39-59.
- ATTIA Mohammed, Younes SAMIH, Ali ELKAHKY *et al.* (2018), « Multilingual multi-class sentiment classification using convolutional neural networks », *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, European Language Resources Association, Miyazaki, <https://aclanthology.org/L18-1101/>, page consultée le 10 octobre 2025.
- BELLI Jill (2016), « Unhappy? There's an app for that: Tracking well-being through the quantified self », *Digital Culture & Society*, 2(1), DOI : <https://doi.org/10.14361/dcs-2016-0107>.
- BERGER Jonah et Katherine MILKMAN (2013), « Emotion and virality: What makes online content go viral? », *NIM Marketing Intelligence Review*, 5(1), p. 18.
- BERNARD Julien (2015), « Les voies d'approche des émotions », *Terrains/Théories*, 2, DOI : <https://doi.org/10/ggwrbb>.

BLAIR-GOLDENSOHN Sasha, Kerry HANNAN, Ryan MCDONALD *et al.* (2008), « Building a sentiment summarizer for local service reviews », *WWW Workshop on NLP in the Information Explosion Era*, p. 339-348.

BURKE Moira et Mike DEVELIN (2016), « Once more with feeling: Supportive responses to social sharing on Facebook », *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 27 février au 2 mars, San Francisco, p. 1462-1474, <https://research.facebook.com/file/373780360921794/once-more-with-feeling-supportive-responses-to-social-sharing-on-facebook.pdf>, page consultée le 8 octobre 2025.

BURKE Moira, Justin CHENG et Bethany DE GANT (2020), « Social comparison and Facebook: Feedback, positivity, and opportunities for comparison », *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 20 au 25 avril, Honolulu, p. 1-13, <https://research.facebook.com/file/3068383903380990/Social-Comparison-and-Facebook-Feedback-Positivity-and-Opportunities-for-Comparison.pdf>, page consultée le 8 octobre 2025.

BURKE Moira, Cameron MARLOW et Thomas LENTO (2010), « Social network activity and social well-being », *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 10 au 15 avril, Atlanta, p. 1909-1912, <https://research.facebook.com/file/205729714945323/social-network-activity-and-social-well-being.pdf>, page consultée le 8 octobre 2025.

CHESBROUGH Henry, Wim VANHAVERBEKE et Joel WEST (2006), *Open Innovation: Researching a new paradigm*, Oxford, Oxford University Press.

CHIGNARD Simon et Louis-David BENYAYER (2015), *Datanomics. Les nouveaux business models des données*, Limoges, FYP Éditions.

CORBIN Alain, Jean-Jacques COURTINE et Georges VIGARELLO (dir.) (2016a), *Histoire des émotions. 1. De l'Antiquité aux Lumières*, Paris, Seuil.

CORBIN Alain, Jean-Jacques COURTINE et Georges VIGARELLO (dir.) (2016b), *Histoire des émotions. 2. Des Lumières à la fin du XIX^e siècle*, Paris, Seuil.

CORBIN Alain, Jean-Jacques COURTINE, Georges VIGARELLO (dir.) (2017), *Histoire des émotions. 2. De la fin du XIX^e siècle à nos jours*, Paris, Seuil.

COULDRY Nick (2019), « Capabilities for what? Developing sen's moral theory for communications research », *Journal of Information Policy*, 9, p. 43-55, DOI : <https://doi.org/10.5325/jinfopoli.9.2019.0043>.

COUNCILL Isaac, Ryan MCDONALD et Leonid VELIKOVICH (2010), « What's great and what's not: learning to classify the scope of negation for improved sentiment analysis », *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, Association for Computational Linguistics, Uppsala, p. 51-59.

CUI Hang, Vibhu MITTAL et Mayur DATAR (2006), « Comparative experiments on sentiment classification for online product reviews », *Association for the Advancement of Artificial Intelligence*, <https://cdn.aaai.org/AAAI/2006/AAAI06-198.pdf>, page consultée le 8 octobre 2025.

DAGIRAL Éric, Séverine DESSAJAN, Tomas LEGON, Olivier MARTIN, Anne-Sylvie PHARABOD et Serge PROULX (2019a), « Faire place aux chiffres dans l'attention à soi », *Réseaux*, 4(216), DOI : <https://doi.org/10/ggjn85>.

DAGIRAL Éric, Christian LICOPPE, Olivier MARTIN et Anne-Sylvie PHARABOD (2019b), « Le quantified self en question(s). Un état des lieux des travaux de sciences sociales consacrés à l'automesure des individus », *Réseaux*, 4(216), DOI : <https://doi.org/10.3917/res.216.0017>.

- DELEUZE Gilles (1988), *Spinoza Philosophie pratique*, Paris, Minuit.
- DELEUZE Gilles et Félix GUATTARI (1980), *Mille plateaux. Capitalisme et schizophrénie 2*, Paris, Minuit.
- DEMSZKY Dorottya, Dana MOVSHOVITZ-ATTIAS, Jeongwoo KO *et al.* (2020), « GoEmotions: A dataset of fine-grained emotions », *arXiv preprint arXiv:2005.00547*, Cornell University, DOI : <https://doi.org/10.48550/arXiv.2005.00547>.
- DWIVEDI-YU Jane et Alon Y. HALEVY (2022), « “That’s so cute!”: The CARE dataset for affective response detection », *arXiv preprint arXiv:2201.11895*, <https://arxiv.org/pdf/2201.11895>, page consultée le 8 octobre 2025.
- EGELMAN Carolyn D., Emerson MURPHY-HILL, Elizabeth KAMMER *et al.* (2020), « Predicting developers’ negative feelings about code review », *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, 23 au 29 mai, Séoul, DOI : <https://doi.org/10.1145/3377811.3380414>.
- EKMAN Paul (1992), « An argument for basic émotions », *Cognition & Emotion*, 6(3-4), p. 169-200.
- EKMAN Paul et Wallace V. FRIESEN (1969), « The repertoire of nonverbal behavior: Categories, origins, usage, and coding », *semiotica*, 1(1), p. 49-98.
- GILROY-WARE Marcus (2017), *Filling the void: Emotion, capitalism and social media*, Londres, Duncan Baird Publishers.
- GROVER Rohan, Josh WIDERA et Mike ANANNY (2025), « Data disaffection: Toward a relational and affective understanding of datafication », *New Media & Society*, 37(2), p. 115-127, DOI : <https://doi.org/10.1177/14614448241310901>.
- HALLINAN Blake (2019), *Like! feelings and friendship in the age of algorithms*, thèse de doctorat, Boulder, University of Colorado.
- HARZING Anne-Will (2007), « Publish or perish », *Harzing.com*, 6 février, <https://harzing.com/resources/publish-or-perish>, page consultée le 9 octobre 2025.
- HICKSON Steven, Nick DUFOUR, Avneesh SUD *et al.* (2019), « Eyemotion: Classifying facial expressions in VR using eye-tracking cameras », *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, p. 1626-1635, DOI : <https://doi.org/10.1109/WACV.2019.00178>.
- HILLIS Ken, Susanna PAASONEN et Michael PETIT (2015), *Networked affect*, Cambridge, MIT Press.
- IZARD Carol (1977), *Human Emotions*, New York, Springer US.
- JAIN Anuj, Debasis PRADHAN, Abishek KUARNR et Shiwangi SINGH (2025), « Self-quantification and consumer well-being: A meta-analytic review », *Psychology & Marketing*, 42(2), p. 514-538, DOI : <https://doi.org/10.1002/mar.22141>.
- JANASIK-HONKELA Nina (2018), « Reclaiming melancholy by emotion tracking? Datafication of emotions in health care and at the workplace », *Open Cultural Studies*, 1(1), p. 549-558, DOI : <https://doi.org/10.1515/culture-2017-0052>.
- JEGOROVA Marija, Stavros PETRIDIS et Maja PANTIC (2023), « SS-VAERR: self-supervised apparent emotional reaction recognition from video », *2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*, Waikoloa Beach, p. 1-8, DOI : [10.1109/FG57933.2023.10042638](https://doi.org/10.1109/FG57933.2023.10042638).
- KARPPI Tero, Lotta KÄHKÖNEN, Mona MANNEVUO *et al.* (2016), « Affective capitalism », *Ephemera: Theory and Politics in Organization*, 16(4), p. 1-13.

KELLEY Patrick Gage, Yongwei YANG, Courtney HELDRETH *et al.* (2021), « Mixture of amazement at the potential of this technology and concern about possible pitfalls: Public sentiment towards AI in 15 countries », *IEEE Data Eng. Bull.*, 44(4), p. 28-46.

KENNEDY Helen (2018), « Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication », *Krisis : Journal for Contemporary Philosophy*, 38(1), <https://krisis.eu/article/view/37184>, page consultée le 27 février 2025.

KRAMER Adam D.I. (2010), « An unobtrusive behavioral model of “gross national happiness” », *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New, York, p. 287-290, DOI : <https://doi.org/10.1145/1753326.1753369>.

KRAMER Adam D.I. (2012), « The spread of emotion via Facebook », *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 5 au 10 mai, Austin, p. 767-770, <https://research.facebook.com/file/2916163195290181/the-spread-of-emotion-via-facebook.pdf>, page consultée le 9 octobre 2025.

KRAMER Adam D.I., Jamie E. GUILLORY et Jeffrey T. HANCOCK (2014), « Experimental evidence of massive-scale emotional contagion through social networks », *Proceedings of the National Academy of Sciences*, 111(24), p. 8788-8790.

KREUK Felix, Adam POLYAK, Jade COPET *et al.* (2021), « Textless speech emotion conversion using discrete and decomposed representation », *arXiv preprint arXiv:2111.07402*, Cornell University, DOI : <https://doi.org/10.48550/arXiv.2111.07402>.

LI Fangtao, Sheng WANG, Shenghua LIU *et al.* (2014), « Suit: A supervised user-item based topic model for sentiment analysis », *Proceedings of the AAAI Conference on Artificial Intelligence*, 27 au 31 juillet, Québec, DOI : <https://doi.org/10.1609/aaai.v28i1.8947>.

LIU Ariel, Victoria Schwanda SOSIK et Khadine SINGH (2018), « Building empathy: Scaling user research for organizational impact », *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 21 au 26 avril, Montréal, DOI : <http://dx.doi.org/10.1145/3170427.3174352>.

LORDON Frédéric (2010), *Capitalisme, désir et servitude : Marx et Spinoza*, Paris, La Fabrique Éditions.

LU Zhiyun, Liangliang CAO, Yu ZHANG *et al.* (2020), « Speech sentiment analysis via pre-trained features from end-to-end ASR models », *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4 au 8 mai, Barcelone, p. 7149-7153, DOI : 10.1109/ICASSP40776.2020.9052937.

MCDONALD Ryan, Kerry HANNAN, Tyler NEYLON *et al.* (2007), « Structured models for fine-to-coarse sentiment analysis », *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, juin, Prague, p. 432-439, <https://aclanthology.org/P07-1055.pdf>, page consultée le 9 octobre 2025.

MCSTAY Andrew (2020), « Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy », *Big Data & Society*, 7(1), DOI : <https://journals.sagepub.com/doi/10.1177/2053951720904386>.

MEJIAS Ulises et Nick COULDRY (2019), *Datafication, Internet Policy Review*, 8(4), <https://www.econstor.eu/handle/10419/214090>, page consultée le 27 février 2025.

MESNIL Grégoire, Tomas MIKOLOV, Marc'Aurelio RANZATO et Yoshua BENGIO. (2014), « Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews », *arXiv preprint arXiv:1412.5335*, DOI : <https://doi.org/10.48550/arXiv.1412.5335>.

METCALF Jacob et Kate CRAWFORD (2016), « Where are human subjects in big data research? The emerging ethics divide », *Big Data & Society*, 3(1), DOI : <https://doi.org/10.1177/2053951716650211>.

- NORMAN Don (2007), *Emotional design: Why we love (or hate) everyday things*, New York, Basic Books.
- MARINOIU Elisabeta, Mihai ZANFIR, Vlad OLARU *et al.* (2018), « 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism », *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, p. 2158-2167, DOI : 10.1109/CVPR.2018.00230.
- PAPACHARISSI Zizi (2015), *Affective publics: Sentiment, technology, and politics*, Oxford, Oxford University Press.
- PAPACHARISSI Zizi et Maria DE FATIMA OLIVEIRA (2012), « Affective news and networked publics: The rhythms of news storytelling on #Egypt », *Journal of Communication*, 62(2), p. 266-282.
- PHARABOD Anne-Sylvie, Véra NICKOLSKI et Fabien GRANJON (2013), « La mise en chiffres de soi : une approche compréhensive des mesures personnelles », *Réseaux*, 1(177), <https://www.cairn.info/revue-reseaux-2013-1-page-97.htm>, page consultée le 27 février 2025.
- PICARD Rosalind W., Elias VYZAS et Jennifer HEALEY (2001), « Toward machine emotional intelligence: Analysis of affective physiological state », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), p. 1175-1191.
- PICARD Rosalind W. (2000), *Affective computing*, Cambridge, MIT Press.
- PORIA Soujanya, Erik CAMBRIA et Alexander GELBUKH (2015), « Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis », *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, septembre, Lisbonne, p. 2539-2544, DOI : 10.18653/v1/D15-1303.
- REYMOND David et Jean-Max NOYER (2022), « Se former aux humanités numériques par les données ouvertes : les chaînes infométriques de Patent2Net (P2N) », *Humanités numériques*, 5, DOI : <https://doi.org/10.4000/revuehn.3014>.
- ROCHER Jean-Maurice (2024), *Le dispositif participatif d'étudiants sur Moments de WeChat : révélateur de micro-relations avec leur environnement universitaire chinois ?*, thèse de doctorat en sciences de l'information et de la communication, Lyon, Université Lyon 2.
- ROGNON Carine, Benjamin STEPHENS-FRIPP, Jess HARTCHER-O'BRIEN *et al.* (2022), « Linking haptic parameters to the emotional space for mediated social touch », *Frontiers in Computer Science*, 4, DOI : <https://doi.org/10.3389/fcomp.2022.826545>.
- RYOKAI Kimiko, Hayes RAFFLE et Robert KOWALSKI (2012), « StoryFaces: Pretend-play with ebooks to support social-emotional storytelling », *Proceedings of the 11th International Conference on Interaction Design and Children*, p. 125-133, DOI : <https://doi.org/10.1145/2307096.2307111>.
- SCISSORS Lauren, Moira BURKE et Steven WENGROVITZ (2016), « What's in a like? Attitudes and behaviors around receiving likes on Facebook », *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 27 février au 2 mars, San Francisco, p. 1501-1510, DOI : <http://dx.doi.org/10.1145/2818048.2820066>.
- SHAFRAN Izhak et Mehryar MOHRI (2005), « A comparison of classifiers for detecting emotion from speech », *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Volume 1*, 23 mars 2025, Philadelphie, p. I/341-I/344.
- SHUKLA Abhinav, Stavros PETRIDIS et Maja PANTIC (2021), « Does visual self-supervision improve learning of speech representations for emotion recognition? », *IEEE Transactions on Affective Computing*, 14(1), p. 406-420.

- SPINOZA Baruch (1677), *L'éthique*, Chicoutimi, Les classiques des sciences sociales, <https://classiques.uqam.ca/classiques/spinoza/ethique/ethique.html>, page consultée le 9 octobre 2025.
- SRNICEK Nick (2017), *Platform capitalism*, New York, John Wiley & Sons.
- STAR Susan Leigh et Anselm STRAUSS (1999), « Layers of silence, arenas of voice: The ecology of visible and invisible work », *Computer Supported Cooperative Work (CSCW)*, 8, p. 9-30.
- STARK Luke et Kate CRAWFORD (2015), « The conservatism of emoji: Work, affect, and communication », *Social Media+ Society*, 1(2), DOI : <https://doi.org/10.1177/2056305115604853>.
- STARK Luke (2020), « The emotive politics of digital mood tracking », *New Media & Society*, 22(11), p. 2039-2057.
- STARK Luke et Jesse HOEY (2021), « The ethics of emotion in artificial intelligence systems », *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, p. 782-793, DOI : <https://doi.org/10.1145/3442188.3445939>.
- STURM WILKERSON Heloisa, Martin J. RIEDL et Kelsey N. WHIPPLE (2021), « Affective affordances: Exploring Facebook reactions as emotional responses to hyperpartisan political news », *Digital Journalism*, 9(8), p. 1040-1061.
- SUN Jennifer J., Ting LIU et Gautam PRASAD (2019), « Gla in mediaeval 2018 emotional impact of movies task », *arXiv preprint arXiv:1911.12361*, DOI : <https://doi.org/10.48550/arXiv.1911.12361>.
- TÄCKSTRÖM Oscar et Ryan MCDONALD (2011a), « Discovering fine-grained sentiment with latent variable structured prediction models », *European Conference on Information Retrieval*, Berlin, Springer, p. 368-374.
- TÄCKSTRÖM Oscar et Ryan MCDONALD (2011b), « Semi-supervised latent variable models for sentence-level sentiment analysis », *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, juin, Portland, Association for Computational Linguistics, p. 569-574.
- TELLIS Gerard J., Deborah J. MACINNIS, Seshadri TIRUNILLAI *et al.* (2019), « What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence », *Journal of Marketing*, 83(4), p. 1-20.
- THOMAS David R. (2006), « A general inductive approach for analyzing qualitative evaluation data », *American Journal of Evaluation*, 27(2), p. 237-246.
- TITOV Ivan et Ryan MCDONALD (2008), « A joint model of text and aspect ratings for sentiment summarization », *Proceedings of ACL-08: HLT*, Columbus, Association for Computational Linguistics, p. 308-316.
- TOMKINS Sylvan (1962), *Affect imagery consciousness. Volume I: The positive affects*, Berlin, Springer Publishing Company.
- VAN ECK Nees Jan et Ludo WALTMAN (2011), « Text mining and visualization using VOSviewer », *ISSI Newsletter*, 7(3), p. 50-54.
- WANG Yi-Chia, Moira BURKE et Robert KRAUT (2016), « Modeling self-disclosure in social networking sites », *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 27 février au 2 mars, San Francisco, p. 74-85, DOI : <http://dx.doi.org/10.1145/2818048.2820010>.

Brevets

US20190197073A1

LIU Eric HC et Charles GORAN (2024), « Methods, systems, and media for personalizing computerized services based on mood and/or behavior information from multiple data sources », U.S. Patent No 12,050,655, 30 juillet 2024.

US20200073485A1

AL-HALAH Ziad, Andrew P. AITKEN, Wenzhe SHI et Alexander GELBUKH (2020), « Emoji prediction and visual sentiment analysis », U.S. Patent Application No 16/561,629, 5 mars 2020.

US20180059885A1

GONNEN Lior et Iddo TAL (2020), « System for and method of accessing and selecting emoticons, content, and mood messages during chat sessions », U.S. Patent No 10,824,297, 3 novembre 2020.

US20180165582A1CHA Meeyoung (2018), « Systems and methods for determining sentiments in conversations in a chat application », U.S. Patent Application No 15/373,415, 14 juin 2018.

CN111443794A

不公告发明人 (2015), « Reading interaction method, device, equipment, server and storage medium », CN Patent No 111443794, 10 mars 2015.

US8977374B1

ECK Douglas et Jay YAGNIK (2015), « Geometric and acoustic joint learning », U.S. Patent No 8,977,374, 10 mars 2015.

CN111767431A

朱延东, 袁泽寰 et 王长虎 (2020), « Method and device for video dubbing », CN Patent No 111767431, 13 octobre 2020.

WO2022256584A1

ABHIJIT C. MEHTA Mateusz Dzwonek, Gloria LIU, David CARVER, David CARR, Sabrina TARDIO, Arpith VARGHESE, Sara BEYKPOUR, Smita Mittal GUPTA et Timothy MARKS (2022), « Labeling messages on a social messaging platform using message response information », W.O. Patent No 2022256584, 8 décembre 2022.

ANNEXES

Annexe A

A1. *TikTok What's Next 2024 Shopping Trend Report*, https://ads.tiktok.com/business/library/TikTok_Whats_Next_Shopping_Trend_Report_2024.pdf, page consultée le 27 février 2025.

A2. *Emotion — the holy grail of customer insight*, https://blog.x.com/en_us/topics/insights/2019/but-how-do-you-really-feel, page consultée le 27 février 2025.

A3. *À propos des réactions à votre publicité*, https://www.facebook.com/business/help/368656903954046?ref=search_new_1, page consultée le 27 février 2025.

A4. *Novos recursos que tornam a criação de vídeos do Instagram Reels ainda mais fácil*, <https://creators.instagram.com/blog/instagram-reels-new-features-creators-audio-sound-effects-stickers-templates>, page consultée le 27 février 2025.

A5. *Edit like a pro with these new features*, https://creators.instagram.com/edit-like-a-pro-5-new-tools?locale=en_US, page consultée le 27 février 2025.

A6. *More ways to share photos on Twitter*, https://blog.x.com/en_us/a/2014/more-ways-to-share-photos-on-twitter, page consultée le 27 février 2025.

A7. *Abbott Philippines. Telling emotion-led brand stories with Facebook and Instagram video ads*, <https://www.facebook.com/business/success/abbott-philippines>, page consultée le 27 février 2025.

A8. *Cinq idées pour préparer votre boutique sur Instagram pour la Saint-Valentin*, <https://business.instagram.com/blog/instagram-prepare-shop-valentines-day-gifts>, page consultée le 27 février 2025.

A9. *De l'émotion à l'action : pourquoi les campagnes à la performance devraient s'inspirer de la Saint-Valentin*, <https://www.facebook.com/business/news/insights/from-heart-to-cart>, page consultée le 27 février 2025.

A10. *Feeling the love this Valentine's Day on TikTok*, <https://newsroom.tiktok.com/en-gb/feeling-the-love-this-valentines-day-on-tiktok>, page consultée le 27 février 2025.

A11. *How Can Businesses Celebrate LGBTQ+ Culture and Community Year Round?*, <https://www.facebook.com/business/news/how-businesses-can-celebrate-lgbtq-pride-all-year-round>, page consultée le 27 février 2025.

A12. *You Belong Here: Celebrating the LGBTQIA+ Community on TikTok*, <https://newsroom.tiktok.com/en-us/you-belong-here-celebrating-the-lgbtqia-community-on-tiktok>, page consultée le 27 février 2025.

A13. *Celebrating #BlackTikTok this October*, <https://newsroom.tiktok.com/en-ie/celebrating-ie-blacktiktok-this-october>, page consultée le 27 février 2025.

A14. *Black Music Month x Pride: Hear From TikTok's Black LGBTQ+ Music Creators*, <https://newsroom.tiktok.com/en-us/black-music-month-x-pride-celebrating-tiktoks-black-lgbtq-music-creators?lang=en>, page consultée le 27 février 2025.

A15. *Unlocking the French-speaking Community*, https://www.tiktok.com/business/library/CA_Canadians_on_TikTok_Unlocking_the_French_speaking_Community.pdf, page consultée le 27 février 2025.

A16. *Découvrez TikTok for Business*, <https://newsroom.tiktok.com/fr-fr/tiktok-for-business-french>, page consultée le 27 février 2025.

A17. *Create a video views campaign*, <https://business.x.com/en/help/campaign-setup/create-a-video-views-campaign>, page consultée le 27 février 2025.

A18. *TikTok Follow Me. Best Practices Playbook*, https://www.tiktok.com/business/library/TikTok_Follow_Me_Best_Practices_Playbook_EN.pdf, page consultée le 27 février 2025.

A19. *Time Well Spent: Users on TikTok stay longer, engage often & feel happier*, https://www.tiktok.com/business/en-US/blog/time-well-spent?ab_version=experiment_1, page consultée le 27 février 2025.

A20. *How Community Generated Entertainment connects communities, culture & brands*, <https://www.tiktok.com/business/en-US/blog/move-aside-ugc>, page consultée le 27 février 2025.

A21. *Meet Symphony Avatars: TikTok's New Creative AI Tool*, <https://newsroom.tiktok.com/en-us/announcing-symphony-avatars>, page consultée le 27 février 2025.

A22. *Furthering our safety and privacy commitments for teens on TikTok*, <https://newsroom.tiktok.com/en-ie/furthering-our-safety-and-privacy-commitments-for-teens-on-tiktok-ire>, page consultée le 27 février 2025.

A23. *Introducing: Shops on Instagram*, <https://business.instagram.com/blog/introducing-shops-on-instagram>, page consultée le 27 février 2025.

A24. *Nos conseils pour concilier les attentes de la clientèle et les incertitudes commerciales pendant la période des fêtes*, <https://www.facebook.com/business/news/balancing-consumer-demands-holiday-tips>, page consultée le 27 février 2025.

A25. *Comment vendre vos produits sur Instagram*, https://business.instagram.com/blog/how-to-sell-your-products-on-instagram?locale=fr_FR, page consultée le 23 septembre 2024.

A26. *Inspiring the Explorers*, https://business.instagram.com/blog/inspiring-the-explorers?locale=en_GB, page consultée le 27 février 2025.

A27. « Emotions drive sales », *The top 10 brands that make people happier on Twitter*, https://blog.x.com/en_us/topics/insights/2019/the-top-10-brands-that-make-people-happier-on-twitter, page consultée le 27 février 2025.

A28. « Joyful discovery leads to spirited shopping », *Holidays for you 2024*, https://www.tiktok.com/business/library/SMB_2024_Holiday_Playbook_AUNZ.pdf, page consultée le 27 février 2025.

A29. *#LoveIsLove — Celebrating LGBTQ Pride Month*, https://blog.x.com/en_us/a/2016/loveislove-celebrating-lgbtq-pride-month, page consultée le 27 février 2025.

A30. *Holidays for you 2024*, https://www.tiktok.com/business/library/CA_Holiday_2024_Vertical_Playbook.pdf, page consultée le 27 février 2025.

A31. *18 trends that highlight fundamental shifts in culture*, https://blog.x.com/en_us/topics/insights/2019/18-trends-that-highlight-fundamental-shifts-in-culture, page consultée le 27 février 2025.

A32. *Watching The Pandemic. What YouTube trends reveal about human needs during COVID-19*, <https://www.youtube.com/trends/articles/covid-impact/>, page consultée le 27 février 2025.

A33. *Celebrating Instagram's Creators of Tomorrow*, <https://creators.instagram.com/blog/instagram-creators-of-tomorrow>, page consultée le 27 février 2025.

A34. *Strengthening how we protect and secure our platform in the US*, <https://newsroom.tiktok.com/en-us/strengthening-how-we-protect-and-secure-our-platform-in-the-us>, page consultée le 27 février 2025.

A35. *Allyship right now: #BlackLivesMatter*, https://blog.x.com/en_us/topics/company/2020/allyship-right-now-black-lives-matter, page consultée le 27 février 2025.

A36. *TikTok puts women's rugby centre stage for the 2023 TikTok Women's Six Nations*, <https://newsroom.tiktok.com/en-gb/womens-six-nations-2023>, page consultée le 27 février 2025.

NOTES

1. Notre traduction de « Online messages influence our experience of emotions, which may affect a variety of offline behaviors ».
2. La BET a été principalement développée par Ekman et Friesen, en 1969. D'autres auteurs comme Tomkins (1962) ou Izard (1977) sont fréquemment associés à cette théorie.
3. Selon certains auteurs, voire entre les articles d'un même auteur, le nombre d'émotions basiques peut varier : la surprise ou le mépris sont parfois intégrés, tantôt rejetés (le premier étant juste un schéma attentionnel, tandis que le second s'exprime avec trop de variations culturelles).
4. Notre traduction de « The *sine qua non* for emotion should not be a unique pan-cultural signal ».
5. Notre traduction de « Cluster individuals into categories ».
6. Notre traduction de « There are of course, many more applications; these are just a sample of what is possible now and in the future ».
7. Voir *Alphabet Announces Fourth Quarter and Fiscal Year 2024 Results*, <https://www.abc.xyz/assets/a3/91/6d1950c148fa84c7d699abe05284/2024q4-alphabet-earnings-release.pdf>, page consultée le 27 février 2025.
8. Voir *Meta Reports Fourth Quarter and Full Year 2024 Results*, https://s21.q4cdn.com/399680738/files/doc_financials/2024/q4/Meta-12-31-2024-Exhibit-99-1-Final.pdf, page consultée le 27 février 2025.
9. Notre traduction de « the requirement, not just the possibility, that every variation in the texture of human experience be translated into data for counting and processing ».
10. Quatrième trimestre 2024 : 13 116 millions chez Alphabet, 12 180 millions chez Meta (*ibid.*).
11. Les brevets sont en effet déposés au nom de Google Inc., et non d'Alphabet.
12. Pour Facebook : <https://www.facebook.com/business>, <https://investor.fb.com>, <https://fr-fr.workplace.com/>. Pour Instagram : <https://business.instagram.com>, <https://creators.instagram.com/>. Pour TikTok : <https://newsroom.tiktok.com>, <https://www.tiktok.com/business>. Pour Twitter/X : <https://business.x.com>, <https://blog.x.com>. Pour YouTube : <https://support.google.com/youtube/>, <https://www.youtube.com/ads/>, <https://blog.youtube/>, <https://www.youtube.com/trends/>, pages consultées le 10 avril 2025.
13. Par exemple, TikTok a créé le néologisme *Newstalgia* (voir annexe A1).
14. Notre traduction de « Emotion can be a mental state that is associated with physiological activity and can be influenced by internal or external environmental conditions ».
15. Tous les verbatims en français et en anglais, comme des liens menant aux pages analysées, sont disponibles dans l'annexe A en fin d'article. Ceux-ci sont référencés par la lettre A et numérotés par ordre d'apparition dans le texte.
16. Notre traduction de « To give people the power to build community and bring the world closer together ».
17. « We aim to fuel the creator economy by investing in creative solutions that spark joy, imagination and action ».
18. Notre traduction de « The last sunny corner of Internet »
19. Notre traduction de « transforming interest into awareness and desire into sales, with a success rate that can be measured ».

20. Ainsi, en 2016, on pouvait lire sur le blogue de Twitter : « We are proud of these efforts and honored to receive a perfect 100 % score from the Human Rights Campaign as a Best Place to Work for LGBT Equality » (A29).

21. Voir : « Meta revoit sa politique de modération de contenus et se rapproche de Donald Trump » par Christophe Dansette sur le site de France 24. En ligne : <https://www.france24.com/fr/%C3%A9missions/info-%C3%A9co/20250108-meta-revoit-sa-politique-de-mod%C3%A9ration-de-contenus-et-se-rapproche-de-donald-trump>, page consultée le 27 février 2025.

RÉSUMÉS

À partir de corpus de publications scientifiques, de brevets et de discours promotionnels en provenance de quatre grandes entreprises propriétaires de plateformes de médias sociaux (Google, Meta, ByteDance et X), nous analysons leurs processus de mise en données des expressions émotionnelles des usagers et usagères. Nous montrons que cette production de données affectives repose sur des représentations universalisantes des émotions et engendre une standardisation de ce qui peut affecter comme une accapARATION commerciale de ce qui émeut en ligne.

From a body of scientific publications, patents, and promotional messages from four major companies that own social media platforms (Google, Meta, ByteDance, and X), we analyze processes for the datafication of users' emotional expressions. We show that the production of affective data is based on generalized representations of emotions and leads to a standardization of what can affect people, as a commercial appropriation of what moves people online. [79 mots]

Con base al corpus de publicaciones científicas, patentes y discursos de promoción procedentes de cuatro grandes empresas propietarias de plataformas de medios de comunicación social (Google, Meta, ByteDance y X), analizamos el proceso de configuración en datos de las expresiones relativas a las emociones de los usuarios. La producción de datos afectivos se basa en representaciones universales de dichas emociones y genera una normalización de lo que puede tener una incidencia debido al acaparamiento comercial de lo que conmueve en línea.

INDEX

Palabras claves : afecto, patentes, dataficação, discurso, emoción, medios de comunicación social

Keywords : affect, patents, datafication, message, emotion, social media

Mots-clés : datafication, affect, émotion, médias sociaux, brevets, discours

AUTEURS

CAMILLE ALLOING

Camille Alloing est professeur à l'Université du Québec à Montréal — Labfluens — LabCMO.
Courriel : alloing.camille@uqam.ca

AMBRE MARIONNEAU

Ambre Marionneau est doctorante à l'Institut national de la recherche scientifique. Courriel : ambre.marionneau@inrs.ca

JULIEN PIERRE

Julien Pierre est professeur agrégé, membre du Département de communication à l'Université de Sherbrooke et Labfluens. Courriel : julien.pierre@usherbrooke.ca

FABIEN RICHERT

Fabien Richert est professeur à l'Université du Québec à Montréal. Courriel : richert.fabien@uqam.ca

Promesses technomnésiques et *datafication* de la mémoire personnelle

Un état des lieux des études critiques sur les données

Rémi Rouge, Emmanuelle Caccamo et Katharina Niemeyer

Pourquoi s'intéresser à la datafication mnésique ?

- ¹ En 2021, Google Photos lançait la fonctionnalité « Little Patterns ». À partir d'un ensemble de photographies stockées dans le téléphone mobile des utilisateur·rices, l'application proposait une sorte de « proto-récit » sous la forme d'un titre (ex. : « Les aventures du sac orange »¹) et d'un album constitué automatiquement par des algorithmes. Plus récemment, par le biais d'une notification thématique (telle qu'« un après-midi ensoleillé à Montréal »), le Journal d'Apple suggère aux individus de rédiger un souvenir au sujet d'un événement récent. Des assistants mnésiques comme Limitless, Omi, Plaud et Memoro² promettent quant à eux d'enregistrer l'environnement sonore et de capturer des conversations tout au long de la journée pour produire une synthèse automatique de l'expérience vécue et faire mémoire. L'agent conversationnel Personal AI va encore plus loin en proposant un « double » numérique de soi qui pourrait raconter à notre place notre parcours de vie, nos souvenirs, et ce, même après la mort. Tirant profit de l'accumulation des données « du temps présent » qui deviennent dans la durée des données du passé dotées d'un potentiel émotionnel fort, ces technologies mnésiques transforment les données³ stockées sur nos nuages et nos appareils en souvenirs automatisés, agencés et fabriqués ; souvenirs qui deviennent à nouveau des formes d'expérience du temps présent. Elles promettent ainsi de créer automatiquement des récits mémoriels personnalisés. Elles le font, d'une part, en exploitant de très importants volumes de données dans des processus dynamiques et récursifs, ce qui invite à poursuivre la réflexion lancée par Beer et Burrows (2013) sur la vitalité de données, trop souvent représentées comme un « matériau dormant » (Denis, 2018). Et elles le font, d'autre

part, à partir de marqueurs ou d'indicateurs singuliers qui reposent sur une interprétation des données, c'est-à-dire en filtrant ce qui « mérite » de devenir souvenir ou non.

- 2 Ce faisant, elles prolongent ce que José Van Dijck appelle l'« idéologie du dataïsme » qui s'appuie sur « la croyance répandue d'une quantification objective et d'un potentiel traçage de tous types de comportement humain » (2014, p. 198). En effet, ainsi saisie par les technologies mnésiques, la mémoire, comme toute expérience humaine, est *datafiée*, c'est-à-dire non seulement *réduite* à un ensemble de données quantifiables pouvant faire l'objet d'un calcul algorithmique, mais également *produite* par des calculs algorithmiques présidant, entre autres, à l'organisation des contenus personnels en ligne et à la sélection des souvenirs qui vaudraient la peine d'être remémorés. Ce n'est donc plus uniquement le stockage des données sous forme de collections ou d'archives qui est en jeu, mais également la façon dont ces données stockées remontent à la surface pour devenir souvenir et, in fine, mémoire.
- 3 Comme l'explique très bien Rik Smit (2024, p. 219), c'est essentiellement aux enjeux économiques et politiques liés aux plateformes numériques que les travaux sur la *datafication* et l'automatisation mnésiques se sont jusqu'à présent consacrés. Cet article explore un aspect inédit de la *datafication* mnésique en se concentrant sur la rhétorique promotionnelle mobilisée par les grands et les petits acteurs pour faire de nos données enregistrées une expérience de mémoire automatisée. Il se focalise plus particulièrement sur les *promesses* liées aux usages des technologies mnésiques émergentes énoncées dans les discours des entreprises qui les produisent. En enquêtant sur ces promesses, cet article renseigne également sur la manière dont les industries du Web envisagent de construire la mémoire personnelle et donc sur les imaginaires de la *datafication* de la mémoire personnelle associés. Rendre compte d'une partie des imaginaires du développement technologique (ici de la *datafication*) peut passer par l'analyse de certains objets du discours ; la promesse constitue une forme et un acte discursifs importants (Joly, 2013).
- 4 Pour ce faire, l'analyse qui suit se situe au croisement de deux champs de recherche. Premièrement, elle s'inscrit dans les approches critiques des *memory studies* et des études en communication et médias. Les *memory studies* constituent un champ interdisciplinaire et international visant une compréhension fine des mécanismes cognitifs, technologiques et culturels de la mémoire (et, de manière trop peu fréquente encore, de l'oubli). Les *memory studies* se penchent aussi sur l'industrialisation et la marchandisation de la mémoire personnelle⁴. Bien qu'il ne s'agisse pas de phénomènes nouveaux (Stiegler, 1991 ; Niemeyer et Douaud, 2018 ; Rouge, 2019), une nouvelle étape a été franchie dans la production de mémoire et de souvenirs à partir de données numériques avec les développements des réseaux socionumériques, d'outils numériques de stockage et d'algorithmes de plus en plus complexes (Rouge, 2023 ; Wang et Hoskins, *op. cit.*) ; ceux mobilisant les technologies de l'intelligence artificielle (IA) en particulier marquent une nouvelle étape sur laquelle reviendra largement cet article. Une partie de la mémoire personnelle se trouve non seulement platformisée (Smit, *op. cit.*), mais également forcée et exploitée (*data mining* mnésique).
- 5 Deuxièmement, sans qu'elle soit au cœur de l'approche choisie, l'article convoque certaines définitions de l'économie de la promesse étudiée par la sociologie des attentes (Dandurand *et al.*, 2022) pour nourrir la réflexion sur la *datafication* mnésique. Au sein des imaginaires sociotechniques, la promesse participe en effet aux attentes que l'on se

fait du développement technologique. Les promesses technoscientifiques contribuent ainsi à façonner les horizons d'attente, elles sont donc performatives et sont surtout porteuses de charge normative (*ibid.*, p. 12-13). C'est à la « dimension discursive » (Joly, *op. cit.*, p. 241) des promesses technoscientifiques que s'attache pour l'essentiel l'analyse⁵ :

[...] la rhétorique de la promesse technoscientifique [forme] un ensemble de moyens d'expression qui présentent la technologie comme une solution inéluctable aux enjeux des sociétés. [...] S'intéresser à la rhétorique de la promesse nous invite à interroger les processus communicationnels employés par les acteurs de l'innovation [...] (Dandurand *et al.*, *op. cit.*, p. 52-53).

- 6 L'article se confronte donc à un registre de promesse spécifique associé à la médiation de la mémoire par des technologies numériques que développent les entreprises étudiées, les promesses *technomnésiques*⁶. Il adopte le double sens du terme *mnésique*⁷, lequel peut se nouer d'abord à la question de la cognition, du fonctionnement cérébral et psychologique de la mémoire, à l'art de se souvenir par le biais des artefacts ou à ses processus involontaires, et encore aux technologies de mémoire ou au stockage de la mémoire sur le plan des objets ou des données.
- 7 Sans présumer de leur efficacité réelle, qu'une étude complémentaire sur la réception de ces promesses par les utilisateur·rice·s pourrait examiner, l'intérêt de faire l'examen critique de ces promesses repose sur le fait que ces discours participent des « injonctions dispositives » (Morelli et Lazar, 2015) qui encadrent les usages de ces dispositifs.
- 8 Ces postulats permettent de formuler les questions de recherche suivantes qui ont guidé l'analyse. Quelles promesses technomnésiques les entreprises étudiées formulent-elles aux utilisateur·rice·s de leurs applications ? Quelles représentations de la mémoire et des données personnelles sont forgées et mobilisées dans ces discours pour permettre la production de cette mémoire numérique ?

Corpus et méthode

- 9 Dans une perspective d'analyse qualitative dans un contexte numérique (Millette *et al.*, 2020), ce sont des documents de communication accessibles en ligne sur les sites Web officiels de quatre entreprises ou projets proposant des technologies mnésiques qui ont été collectés (par capture d'écran ou téléchargement) : Google⁸, Apple, Memoro et Personal AI (Human IA Labs)⁹. Pour les deux premiers, la fonctionnalité des photographies est celle qui est le plus associée aux souvenirs, mais pour Apple l'application visionOS2 joue également un rôle important.

Google Photos, fonctionnalité « Souvenirs »

- 10 L'ensemble des articles consacrés au produit Google Photos sur le site de l'entreprise¹⁰ ont été consultés, c'est-à-dire 57 articles publiés depuis le lancement officiel du produit le 28 mai 2015 au dernier article publié lors de la rédaction de cet article, le 6 décembre 2024. Une quinzaine d'articles, plus directement en lien avec les thématiques de l'enquête, ont fait l'objet d'une analyse plus approfondie.

Figure 1. Capture d'écran de la page consacrée à Google Photos

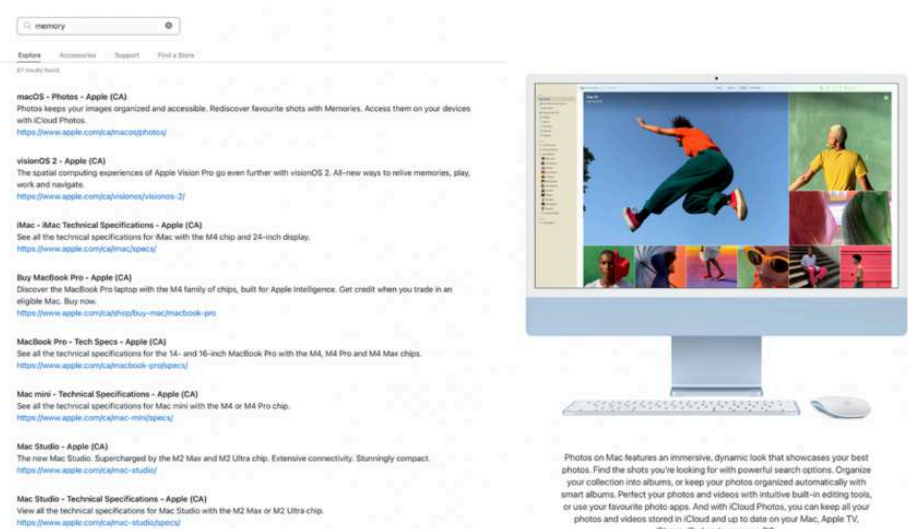


© Google, février 2025, <https://blog.google/products/photos/>

Apple Photos, fonctionnalité « Apple Souvenir » et visionOS 2

- 11 En vue d'accéder à l'ensemble des publications qui mentionnent le mot-clé *mémoire*, la requête de recherche avec le mot-clé « *memory* » a été réalisée sur le site Web d'Apple : 87 pages s'affichent et les deux premières (fig. 2) sont Apple Photos et visionOS2 (expériences augmentées), suivies de références aux outils techniques ou augmentation d'espace de mémoire et de stockage. Les deux premières pages Web ont donc été consultées et téléchargées en détail, car elles concernent principalement la problématique de cette recherche, mais certains hyperliens d'Apple Photos renvoient vers des fonctionnalités ou besoins d'espace de stockage et sont mentionnés dans la présentation des résultats. Des captures d'écran ont été réalisées pour garder trace de l'arrangement de la page et de l'ordre des énoncés et des images (fig. 2).

Figure 2. Capture d'écran de la requête « memory » avec exemples



© Apple, janvier 2025, <https://www.apple.com/ca/macos/photos/>

Memoro (depuis 2024, MIT)

- 12 Le corpus rassemblé comprend la page de présentation du projet du Massachusetts Institute of Technology (MIT)¹¹ depuis son lancement en 2024, y compris les textes, visuels, vidéos et un article savant coécrit par l'inventeur de Memoro.

Personal AI (depuis 2021, Human AI Labs)

- 13 Les pages de présentation du projet sur le site Web de Personal AI ont été consultées depuis le lancement en 2021, en particulier la « thèse » de l'entreprise¹², sa description de la mémoire¹³ et la documentation fournie¹⁴. L'application a aussi été testée.
- 14 L'analyse porte donc à la fois sur un corpus d'articles ou contenus de page Web promotionnels issus des sites Web des technologies concernées et également sur la mise en scène visuelle, formelle et procédurale (interactive) des interfaces de ces différentes technologies.
- 15 Le corpus est assez hétérogène et ne peut donc pas être facilement comparable sur le plan sémiotique. Il a donc d'abord été question de tenir compte du vocabulaire mobilisé en lien avec la *datafication* mnésique et de comparer les entreprises. L'ancrage épistémologique de l'analyse qualitative adoptée est celui de la sociosémiotique¹⁵ ouverte (Pailhes et Sapio, 2022), qui permet d'explorer et de saisir dans les discours multimodaux (textuels, visuels, audiovisuels, procéduraux) l'univers social, économique et culturel de la production et de la réception de ces technologies.
- 16 La première partie de l'article s'intéresse à la promesse générale d'enregistrement et d'automatisation de nos expériences quotidiennes sous la forme de données numériques qui deviennent in fine, par leur stockage et leur réactivation potentielle, matière de mémoire pour une production mnésique. L'analyse des discours des entreprises étudiées a par ailleurs permis d'identifier des « sous-promesses » nouées à

cette promesse générale, étudiées dans la seconde partie de l'article : celle d'une mémoire « augmentée » associée, d'une part, à la classification des données, d'autre part, à la production de « souvenirs numériques », et enfin, celle d'un redoublement durable de l'expérience afin de « lutter » contre l'oubli et la disparition du sujet ou de la personne.

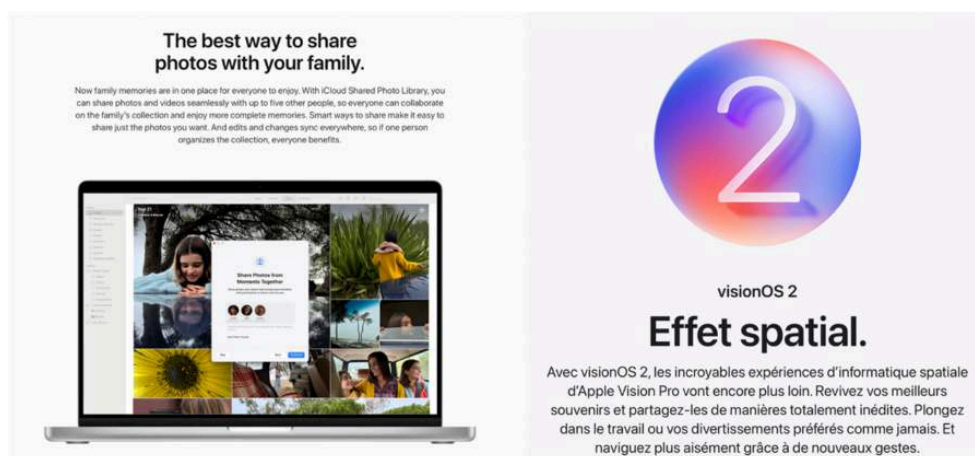
Une promesse d'automatisation : des « problèmes », des « solutions »

- 17 Malgré la diversité des technologies proposées, la construction des attentes par les entreprises étudiées repose invariablement sur une promesse d'automatisation des processus mnésiques. Pour garantir l'utilisation des technologies ou des applications qu'elles commercialisent, ces entreprises produisent des discours mettant en avant leur caractère incontournable, qu'elles adossent à une vision du monde étroitement liée à l'idéologie du progrès et du dataïsme. Autrement dit, les technologies mnésiques émergentes sont présentées comme des outils essentiels et nécessaires à la gestion de la mémoire personnelle sans lesquelles certains souvenirs seraient voués à disparaître.
- 18 La promesse d'une gestion automatique des données de la mémoire personnelle s'appuie sur un constat plus ou moins tacite : avec les évolutions récentes, dans le domaine de l'IA notamment, une nouvelle étape de l'histoire de l'externalisation de la mémoire se présenterait¹⁶, étape caractérisée par la réduction potentielle de ce qui constitue la mémoire humaine à un ensemble de données quantifiables et traitables au moyen d'outils algorithmiques (Hoskins, *op. cit.*).
- 19 La construction de cette promesse repose avant tout sur une « première étape », celle de produire des « problèmes » à résoudre (Joly, *op. cit.*, p. 236). Les « problèmes », et le rôle joué par la *datafication*, prennent deux grandes formes différentes.
- 20 Le premier problème formulé est celui du trop-plein de données auquel feraient face les individus dans un monde numérisé. De ce point de vue, les technologies mnésiques constitueraient autant de « solutions » à ce problème. En d'autres termes, il s'agit de chercher ce qui pourrait répondre à des « besoins » des utilisateur·rice·s¹⁷ pour produire de nouvelles fonctionnalités (ou alors de trouver à quel besoin une fonctionnalité qu'on souhaite développer peut répondre) afin de présenter le produit comme réponse nécessaire à des attentes légitimes¹⁸.
- 21 C'est sur cet argument que repose la stratégie de légitimation de plusieurs produits étudiés. Les articles consacrés à Google Photos sur le site de l'entreprise¹⁹ mettent très souvent en avant le trop grand nombre de photos et vidéos dont nous disposons et notre difficulté à les gérer. Cette justification apparaît dès l'article annonçant le lancement du produit :
[...] the more moments we capture, the more challenging it becomes to relive those memories. Photos and videos become littered across mobile devices, old computers, hard drives and online services (which are constantly running out of space). [...] We wanted to do better. So today we're introducing Google Photos — a new, standalone product that gives you a home for all your photos and videos, helps you organize and bring your moments to life, and lets you share and save what matters²⁰.
- 22 On la retrouve à diverses reprises, soit pour justifier a posteriori le choix de développer Google Photos²¹, soit pour légitimer le développement de nouvelles fonctionnalités,

comme c'est le cas dans le dernier article mettant en avant l'intérêt de coupler Google Photos avec le robot conversationnel de l'entreprise, Gemini : « Often we snap photos to capture memories or save important details, but the information in those photos can get lost in the shuffle. [...] Gemini's multimodal capabilities can understand the context and subject of photos to pull out details²². »

- 23 La rhétorique de la réponse à un problème se retrouve aussi dans la présentation du service Photos d'Apple et de visionOS2 (fig. 3). Par exemple, lorsque la survivance de la beauté du souvenir familial ou encore la promesse de revivre les souvenirs sont annoncées, les slogans sont rapidement accompagnés d'une autre annonce portant sur la nécessité d'augmenter la mémoire de stockage ou d'acheter des nouveaux accessoires permettant d'accéder aux nouvelles fonctionnalités²³ ; une sorte de double injonction jonglant entre accumulation de données et promesse de solution pour ces dernières se met en place. Ou encore, lorsqu'il s'agit de justifier la fonctionnalité permettant le partage des souvenirs de famille en un seul lieu par la iCloud Shared Photo Library. Censées répondre au problème de la décentralisation des données, les bibliothèques de photos s'unifient pour devenir mémoire de groupe, ce qui nécessite l'installation des applications et donc la possession des produits de la compagnie pour plusieurs personnes.

Figure 3. Captures d'écran d'Apple Photos et de visionOS2



- 24 La particularité de cette manière de construire la légitimité du produit en le présentant comme la réponse à un problème est que les deux termes, le problème et la solution, sont de même nature : ils procèdent tous deux de la *datafication* de nos expériences quotidiennes²⁴. Plus significatif encore, la solution au problème participe à l'entretenir, ou du moins à le reconduire. L'usage de produits comme Google Photos ou Apple Photos permet aux utilisateur·rice·s de continuer non seulement à accumuler toujours plus de contenus, mais aussi à les maintenir, de ce fait, dans une forme de dépendance à ces solutions technologiques, comme l'illustre la citation suivante, issue du blogue de Google en 2015 : « As you keep snapping photos and capturing videos, we'll keep working on making them even easier to store, share and bring to life²⁵. »
- 25 L'autre grand « problème » auquel les technologies sont censées répondre est celui du caractère faillible de la mémoire humaine, qui est perçue comme défaillante dans son fonctionnement. On trouve cette idée explicitement formulée dans le discours de Personal AI en 2022 (Human IA Labs) : « At its core, memory retention is a biological

problem. But instead of focusing on fixing people with brain enhancement drugs or surgical implants, why not build a tool to solve the problem²⁶? »

- 26 Les technologies mnésiques viendraient répondre à ce « problème » de la mémoire oublieuse. À cela s'ajoute la disparition potentielle de la mémoire personnelle, notamment après la mort du sujet. Le passage du temps effacerait les souvenirs, et avec eux la mémoire des individus²⁷. Dans ce récit, la technologie est présentée comme la « solution » à un problème quasi anthropologique, comme on peut l'observer dans l'explication fournie à l'onglet Memory du site de Personal AI : « For all of human history, a person's memory could be lost; now, in the digital age — and more specifically, in the AI age — your memory can be recalled and shared seamlessly²⁸. »

- 27 Ce discours de la lutte contre l'oubli, voire d'un refus de l'oubli pourtant essentiel à la mémoire personnelle (Ricœur, 2003 ; Todorov, 2004 ; Wang et Hoskins, *op. cit.*), peut prendre une forme moins spectaculaire lorsque la technologie est présentée comme un outil pour pallier les « défaillances » normales de la mémoire de l'utilisateur-riche dans des situations plus ordinaires, comme le propose par exemple Memoro. Grâce à un enregistrement continu de l'environnement sonore des utilisateur-riche-s, cette technologie va inférer, voire anticiper des « besoins » mnésiques en temps réel²⁹. La technologie est décrite comme un assistant personnel de mémoire, doté d'une IA, « wearable » avec interface vocale. Il s'agit d'un assistant du quotidien qui promet de devancer/anticiper les besoins des utilisateur-riche-s lorsqu'ils et elles ne se souviennent plus du nom de la personne devant eux, de certains détails ou de la liste des produits à acheter à l'épicerie. Il s'agit donc de pallier les défaillances de la mémoire dite « sémantique », qui renvoie notamment aux connaissances apprises, aux informations générales sur des lieux et les personnes. Memoro souffle vocalement les réponses à l'oreille de l'utilisateur-riche qui n'aurait pas retenu ce type d'information au cours de la journée :

People have to remember an ever-expanding volume of information. To help with this, we developed Memoro, a wearable and audio-based personal memory assistant that uses an AI agent. Memoro automatically captures real-world auditory interactions of the user and converts them into digital memories. During recall, the assistant can directly infer memory needs using context, search, and present minimal suggestions to weave into daily life and conversations³⁰.

- 28 Il est intéressant de noter que, dans son discours promotionnel, Memoro joue sur une frontière entre le médical et l'extramédical : la technologie peut à la fois aider des personnes âgées avec perte de mémoire et toute la population qui voudrait bénéficier d'une mémoire plus « performante »³¹.
- 29 Dans la communication de Google Photos, bien que le thème de l'oubli soit moins fréquent, il est tout de même mobilisé, par exemple pour justifier le développement d'une nouvelle fonctionnalité, « Suggested sharing », qui permet de pallier l'oubli de partage des souvenirs avec nos proches³².
- 30 Dans tous ces cas, la *datafication* et le traitement algorithmique des données sont présentés comme une solution incontournable à un « problème », qu'il soit d'ordre technique (la numérisation du monde entraînerait la production toujours croissante de données difficiles à gérer sans assistance technologique) ou anthropologique (la mémoire « défaillante », l'oubli comme corrélat nécessaire de la mémoire), qui consiste à prétendre automatiser des processus mnésiques.

- 31 En fonction de ce qui fait l'objet de l'automatisation et des « problèmes » auxquels elle prétend répondre — autrement dit, en fonction de comment et pourquoi les données produites sont mobilisées —, différentes sous-promesses découlent de cette promesse générale d'automatisation. Chacune d'elles peut être perçue comme une prétention à répondre à ces soi-disant « problèmes ».

Les sous-promesses de la *datafication* de la mémoire

- 32 S'appuyant sur ces discours de légitimation, la promesse de l'automatisation de la mémoire peut prendre différentes formes prétendant se substituer à différentes opérations mémorielles qui témoignent de la variété des technologies mnésiques existantes. Dans la suite de notre réflexion, nous explorerons comment la promesse de l'automatisation s'accompagne de l'idée d'une « augmentation » mémorielle, qui se manifeste sous deux formes différentes : l'*anamnésis* et la *mnémé*. L'article explorera enfin la sous-promesse du redoublement de l'expérience quotidienne pour produire de la mémoire et « augmenter » les personnes.

« Augmenter » la mémoire

- 33 L'idée de « mémoire augmentée » par la technique est loin d'être récente³³, mais les technologies à l'étude formulent la promesse de nouvelles formes de mémoire augmentée.
- 34 Cette « augmentation » de la mémoire renvoie en réalité à deux phénomènes mnésiques distincts : un *travail actif de remémoration* et l'*évoation du souvenir* (au sens de contenu réminiscent). Cette distinction en recouvre une bien plus ancienne, que l'on trouve chez Aristote, qui oppose l'*anamnésis* et la *mnémé*, et que Paul Ricœur analyse longuement dans *La mémoire, l'histoire, l'oubli* (2003). Selon la lecture qu'en propose Ricœur, Aristote distingue deux phénomènes mnésiques. L'*anamnésis* renvoie au travail de remémoration qui demande une recherche active de la part d'un sujet qui parcourt une séquence de temps déterminée selon un mouvement choisi par ses soins. L'autre phénomène, appelé la *mnémé*, consiste en la survenance d'un souvenir sous l'effet d'une affection, c'est-à-dire d'un changement de la sensibilité lié à une cause extérieure. Le souvenir survient à l'esprit du sujet presque « passivement » (*ibid.*, p. 4). À travers leurs différentes fonctionnalités, les technologies étudiées proposent d'augmenter la mémoire en prenant en charge l'un, l'autre ou, le plus souvent, les deux phénomènes.

L'*anamnésis* algorithmisée

- 35 Disponibles dans les téléphones cellulaires ou dans d'autres types d'appareils à porter sur soi, les fonctionnalités/technologies mnésiques se présentent comme des « lieux de mémoire en mouvement » (Niemeyer, 2011, p. 315) incontournables du *xxi*^e siècle. Grâce à des IA prétendant reproduire nos processus cognitifs, elles promettent de gérer nos souvenirs à notre place. « Your memories, automatically organized³⁴ », telle est la promesse formulée, pour Google Photos, par le *product manager* Ben Greenwood en 2019³⁵. Cette automatisation est censée faciliter un travail d'organisation des contenus en vue de leur remémoration présentée comme laborieuse : « Smarter photo albums,

without the work³⁶ », promettait quant à lui François de Halleux, également *product manager* pour l'entreprise de Mountain View, en 2016.

- 36 Cette « curation » algorithmique de la mémoire personnelle repose sur un ensemble de choix plus ou moins explicitement énoncés par les entreprises. Par la constitution de vastes fonds de contenus à potentiel mnésique, les technologies proposent de simplifier la navigation dans la masse des contenus enregistrés. Préalablement indexés selon une variété de critères de plus en plus nombreux, les contenus personnels des utilisateur·rice·s sont présentés comme facilement accessibles et retrouvables, la numérisation des contenus favorisant leur conservation et simplifiant leur « recherchabilité » (*searchability*). Étroitement liée, comme le souligne Ben Jacobsen dans sa thèse (2020, p. 120), au développement de l'apprentissage machinique et à la nécessité de produire des outils permettant de naviguer dans les archives des utilisateur·rice·s, la recherchabilité des contenus est un élément très souvent mis en avant dans la communication des entreprises produisant des technologies mnésiques. On en trouve de nombreux exemples dans les articles consacrés à Google Photos sur le site de l'entreprise, par exemple : « Google Photos was one of the first products we built with AI at the center, giving you the ability to search your photos and videos for people, pets, places and more³⁷. » Cette sous-promesse est très souvent associée à un argument consistant à insister sur le gain de temps associé à cette recherchabilité, que souligne également Jacobsen (*op. cit.*, p. 121).
- 37 Les modes d'organisation et de recherche des contenus sont ce faisant construits en fonction des capacités des technologies qu'utilisent ces entreprises (Rouge, à paraître), et non pas des besoins ou des repères mobilisés par les utilisateur·rice·s pour reconstituer leurs souvenirs. Si le traitement des métadonnées associées aux photos et vidéos³⁸ rend possibles l'organisation et la recherchabilité des contenus en fonction de leur date et du lieu de capture³⁹, l'organisation des contenus en fonction de différents types de repères visuels (visages, objets, activités, etc.) a été rendue possible par l'utilisation d'algorithmes de reconnaissance visuelle de plus en plus sophistiqués développés par les entreprises produisant ces technologies mnésiques⁴⁰. Alors que les *chatbots* connaissent un succès médiatique considérable, c'est désormais la recherche assistée par des robots conversationnels qui est mise en avant par Google Photos⁴¹.
- 38 Ainsi présentée, tout se passe comme si la *datafication* de la mémoire entraînait une amélioration notable de l'expérience anamnésique des utilisateur·rice·s. C'est toutefois sans tenir compte de la nécessaire sélection préalable de ce qui constitue la mémoire de l'utilisateur·rice en fonction des propriétés et du fonctionnement des différentes technologies étudiées. En effet, en fonction du type de technologies étudiées ou du dispositif dans lequel elles s'inscrivent, la capture des contenus et leur indexation varient, car elles ne prennent pas nécessairement en charge tous les formats de contenus numériques : par exemple, Google Photos ne prendra pas forcément en charge des documents essentiellement textuels, l'entreprise associant ainsi souvenirs et images.

La mnémé algorithmisée

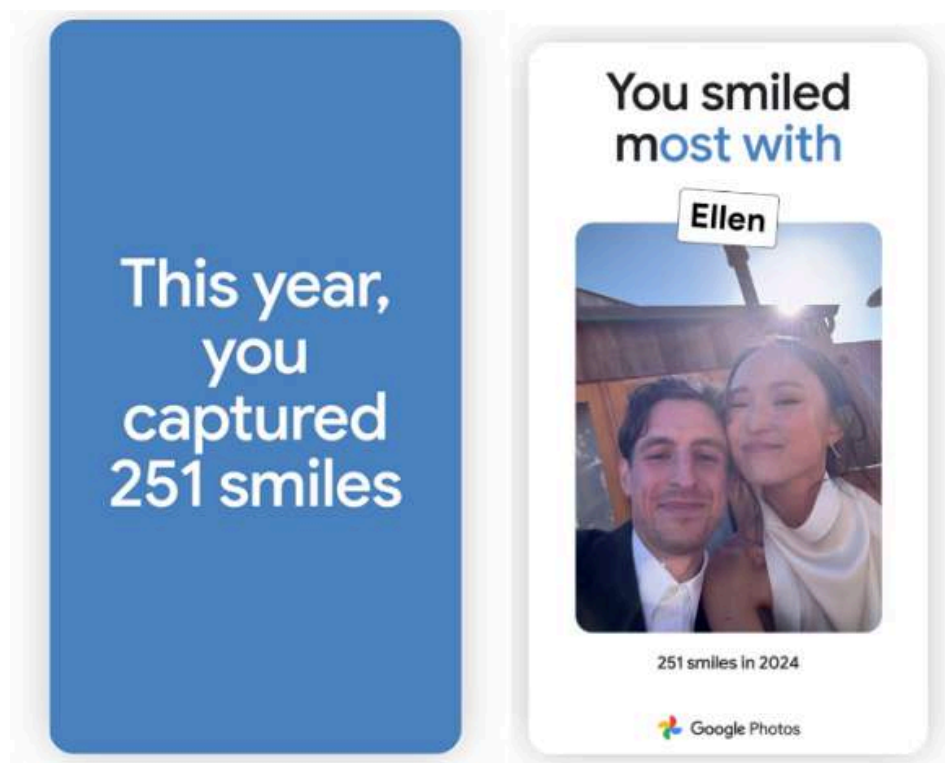
- 39 C'est dans un autre sous-ensemble de discours mettant en avant une sorte d'automatisation de la réminiscence que s'inscrit la promotion des souvenirs numériques, c'est-à-dire des contenus produits par des technologies proposant de

rappeler aux individus des événements particuliers de leur expérience quotidienne (relevant donc de la mémoire « épisodique », celle de moments vécus par l'utilisateur-riche).

- 40 Les technologies algorithmiques jouent ici le rôle d'évocateur de souvenirs : les algorithmes sélectionnent, à partir d'un ensemble d'instructions⁴², des données dignes de revenir dans l'actualité des utilisateur-riche-s. C'est cette fois-ci la survenue du souvenir qui répond au problème de la surabondance de contenus, comme en témoignent différentes formules provenant du blogue de Google Photos, par exemple en 2019 : « You lose the warm and fuzzy nostalgic feeling when you have to scroll through hundreds of duplicate photos, so we're putting your memories front and center in Google Photos. » Ou en 2021 : « To make it easier to look back, we're using AI to power new features that resurface meaningful moments and bring your memories to life [...] »⁴³.
- 41 Bien que les entreprises donnent peu d'informations sur les critères qui président à la sélection des « souvenirs numériques », les interfaces ainsi que les textes publiés, par exemple sur le blogue de Google Photos, permettent de réunir quelques renseignements utiles à ce sujet.
- 42 Le premier élément qui se dégage de plusieurs billets publiés sur le blogue de Google ou sur les pages concernant Apple Photos est que les souvenirs sont de nature photographique et vidéographique, ce qui exclut a priori les contenus de nature plus textuelle. Par exemple, un ticket de caisse, un billet de cinéma ou une recette, pourtant pris en photos et classés comme tels, formeront difficilement des « souvenirs » valables pour être mis en avant par Apple Photos.
- 43 Les captures d'écran sont aussi souvent prises comme exemple de contenus qui ne sont pas dignes d'être rappelés à l'attention de l'utilisateur-riche :
- Our phones are filled with a growing number of photos and videos capturing everything from everyday moments to life milestones. But they're also cluttered with duplicate images, screenshots, and blurry videos, making it hard to focus on the most meaningful moments⁴⁴.
- 44 La capture d'écran est renvoyée à une dimension pratique (« for future reference ») qui l'oppose à des souvenirs caractérisés par leur importance et leur dimension affective : « Google Photos is home for your memories. But we know your photo gallery can get cluttered with similar photos of a single moment or screenshots and documents you snapped for future reference »⁴⁵.
- 45 De plus, la nature vidéo/photographique des souvenirs de Google Photos et d'Apple Photos implique des critères de sélection basés sur la qualité plastique des contenus. Les photos floues, similaires, comprenant des sujets fermant les yeux sont également jugées par Google comme n'étant pas dignes d'être rappelées à la mémoire de l'utilisateur-riche, au point que l'entreprise propose depuis 2024 une fonctionnalité permettant de combiner plusieurs clichés pour obtenir, de manière factice, la « bonne prise » (« Best Take », du nom de la fonctionnalité en question)⁴⁶.
- 46 S'appuyant sur des technologies de reconnaissance faciale, les algorithmes de Google Photos privilégient non seulement les photos de bonne qualité, mais également celles sur lesquelles les sujets sourient : « We're using machine learning to curate what appears in Memories, so you don't have to parse through many duplicate shots, and you can instead reflect on the best ones, where the photos have good quality and all the

kids are smiling⁴⁷. » La rétrospective de l'année 2024 proposée par l'entreprise fait la part belle à ce critère, qui devient même une mesure de la qualité des expériences vécues et des relations entretenues (voir figure 4).

Figure 4. Capture d'écran de Google



© Google, 2024, <https://blog.google/products/photos/google-photos-2024-recap/>

- 47 Ce critère repose sur la conviction qu'ont acquise les professionnel·le·s de ces entreprises selon laquelle il est préférable et possible de sélectionner les contenus susceptibles de susciter des émotions positives. Autrement dit, un « bon souvenir », pour Google Photos, correspond à un souvenir positif (Rouge, 2023 et à paraître).
- 48 Par ailleurs, en mobilisant les métadonnées associées aux photos (Lee, *op. cit.* ; Rouge, 2023), les algorithmes produits par Google Photos sélectionnent les contenus capturés à des endroits peu visités par l'utilisateur·rice, qui constituent de probables lieux de vacances : « The holidays are always a great time to reflect on what really matters — for me, that's my family, both near and back home in Europe. With Memories in Google Photos, I'm able to relive the moments — big and small — I've shared with my loved ones⁴⁸. » Ces métadonnées peuvent également permettre de reconnaître des dates spéciales, correspondant à des événements culturels importants pour les différent·e·s utilisateur·rice·s⁴⁹.
- 49 Enfin, les technologies d'IA sont utilisées pour produire des assemblages de contenus thématiques censés représenter des souvenirs associés à un type d'activité, à un objet, à un lieu. Les algorithmes sont ici instruits pour repérer automatiquement des motifs réguliers dans les contenus, même les plus banals, comme la présence d'un sac à dos orange sur une multiplicité de photos :

With Memories, you can already look back on important photos from years past, recent highlights, moments with your loved ones, your favorite activities and more.

Using machine learning, we can now go beyond resurfacing photos based on themes to doing so based on not-so-obvious visual patterns in your photos. Starting later this summer, when we find a set of three or more photos that share things like shape or color, we'll highlight these little patterns for you in your Memories. For example, one of our engineers received this collection featuring photos he snapped of his favorite orange backpack⁵⁰.

- 50 Apple Photos fonctionne de manière relativement similaire. Pour devenir « souvenirs », les images doivent avoir été classifiées selon un ensemble de catégories prédéfinies : prénoms de personnes, groupe de personnes, animaux (« amis à quatre pattes »), lieux géolocalisés (à domicile, boutiques ou cafés, villes, pays), vacances, saisons et événements calendaires (été, hiver, Noël, veille du Jour de l'an, etc.) ou inférés (« premiers moments ensemble »), scènes (« journées enneigées », « en nature »), actions (« exploration »). Or, certaines classes ou sous-classes ne pourront pas faire l'objet de souvenirs thématiques, comme c'était par exemple le cas avec Apple Photos concernant les lieux de commémoration de l'Holocauste : l'algorithme bloquait les photographies de lieux sensibles de ses rétrospectives annuelles ou des images mises en avant (Vilvang, 2024).
- 51 L'ensemble de ces critères participent, selon les communications promotionnelles d'Apple Photos et de Google Photos, à la production de souvenirs « importants », « significatifs »⁵¹ — c'est-à-dire à rappeler des souvenirs qui comptent⁵². Cette importance des souvenirs est associée chez Google à une affectivité supposée forte que dénote le recours aux superlatifs (« best moments », « best memories⁵³ ») et aux préférences fortes (« your favorite people, things and activities⁵⁴ »). Bien sûr, les suggestions mnésiques proposées peuvent être en total décalage avec l'expérience vécue et les phénomènes de réminiscence des utilisateur·rice·s⁵⁵.
- 52 La valeur des souvenirs — leur poids et leur sens dans l'expérience de l'utilisateur·rice — est donc calculée par les algorithmes de ces entreprises qui opèrent à partir de critères normatifs davantage pensés en fonction des capacités des technologies mobilisées que de critères établis par les utilisateur·rice·s. Ce biais technologique et son emprise sur le récit de vie des utilisateur·rice·s sont par ailleurs verbalisés par la vice-présidente de Google Photos dans le billet de 2021 plusieurs fois cité, lorsqu'elle annonce le lancement des « little patterns », ces assemblages de photos basés sur le repérage de motifs réguliers dans les contenus : « On their own, these photos may not be the most meaningful, but when you see them all together, they tell a story — your story⁵⁶. »
- 53 Le discours d'Apple Photos présente quant à lui une promesse tout à faire inédite dans le domaine de la mémoire personnelle : Apple Souvenir promet une mémoire « équitable » (« fairness »)⁵⁷ aux utilisateur·rice·s. Pour l'entreprise, la mise en avant des « souvenirs » doit prendre en considération des catégories intersectionnelles (ethnoraciales, d'âge et de genre) et la modélisation algorithmique doit être « diversifiée, inclusive et équilibrée⁵⁸ ». L'idée d'associer la *mémoire personnelle* à un critère d'équité forme une orientation idéologique et un cadrage rhétorique forts tout à fait nouveaux. La dimension équitable (ou non équitable) est en effet généralement liée aux questions de mémoire collective (mémoires nationales, des groupes, voire des familles, par exemple), pas d'emblée à la mémoire individuelle. La mémoire personnelle n'est pas présidée par un critère d'équité, contrairement, par exemple, aux politiques de mémoire collectives contemporaines plus inclusives ou aux politiques de réparation mémorielles des États visant à réparer le tort historique

causé à certains groupes sociaux. Cette promesse de rendre la présentation des souvenirs personnels équitable selon les catégories sociales forme de toute évidence un argument commercial et un positionnement inévitable à l'égard des critiques sur le biais des algorithmes.

- 54 En proposant des technologies sélectionnant les contenus dignes d'être rappelés, ces entreprises organisent non seulement une « partition du mémorable » (Jacobson et Beer, 2021, p. 40), mais produisent également une fonctionnalité permettant de justifier, aux yeux des utilisateur·rice·s, l'intérêt de la collecte de données personnelles auxquelles ils et elles se livrent. Cette justification peut prendre des formes différentes qui, toutes, insistent sur la qualité des souvenirs produits, qu'ils soient « significatifs », « valables » ou encore « équitables ». Un effet commun se dégage de ces différentes justifications cependant : les critères de la réminiscence, reposant sur des technologies qui capturent et « interprètent » les données, sont rationalisés. Ils ne dépendent plus d'une construction sociale partagée, d'aléas de l'expérience quotidienne ou du caprice des sens, mais d'un ensemble d'opérations algorithmiques.
- 55 On notera par ailleurs que la sélection automatique du mémorable est aussi souvent associée à celle d'une « augmentation » multimodale de l'expérience de la réminiscence algorithmisée. Des technologies telles que Google Photos ou visionOS2 proposent d'augmenter les souvenirs avec des effets visuels et sonores censés rendre plus immersive l'expérience de la redécouverte de ces contenus : « Cinematic photos help you relive your memories in a way that feels more vivid and realistic — so you feel like you're transported back to that moment⁵⁹ » ou encore « D'un simple geste, vos souvenirs prennent vie, avec un effet de profondeur et de relief infiniment naturel⁶⁰ ». VisionOS2 formule également la promesse de « répéter » à l'infini les expériences vécues et dans n'importe quel endroit au moyen d'un casque de réalité virtuelle : « Pendant vos appels FaceTime avec des personnes équipées d'Apple Vision Pro, tout le monde peut partager des images spatiales grandeur nature — photos, vidéos et panoramas — et interagir avec elles. Vous pouvez ainsi revivre vos plus beaux souvenirs, où que vous soyez⁶¹. » Ces images augmentées deviennent à nouveau des expériences du temps présent tout en mettant l'accent sur l'enrichissement des expériences en devenir. Le souvenir qui ressurgit par le biais de cette *mnémé* algorithmique est ainsi nouvellement augmenté, à la fois par les algorithmes et par un ensemble d'effets multimodaux ajoutés⁶².
- 56 En somme, la distinction entre l'*anamnésis* et la *mnémé* est utile à la compréhension de la différence qui distingue les phénomènes mnésiques que les technologies se proposent d'augmenter : d'un côté, des technologies promettant une organisation et une navigation automatisées dans les archives numériques, relevant d'une sorte d'*anamnèse*, de l'autre (et souvent associées aux premières au sein d'un même produit), des technologies proposant d'automatiser la sélection des contenus mémorables en produisant une suggestion de souvenirs numériques. La publicité d'Apple pour l'iPhone 7, la fonctionnalité « Souvenirs » plus spécifiquement, est possiblement la plus exemplaire de cette double rhétorique de la *mnémé* et de l'*anamnésis*. La dernière est gérée par Apple (un vieil homme qui travaille dans les archives analogiques qui coupe et prépare les données pour leur projection) afin de produire des *mnémés* chargées fortement en émotions chez les personnes qui utilisent Memories⁶³.

« Augmenter » la mémoire et l'utilisateur·rice par une *datafication* en continu

- 57 Automatiser l'enregistrement d'un vaste ensemble de traces en vue de faire mémoire est lié à un autre ensemble de promesses que l'on peut rattacher au phénomène du *lifelogging*. Il s'agit d'abord de produire un prétendu « double » *datafié* de l'expérience vécue par un individu. Idéalement, dans le but de constituer une sorte de « mémoire totale » du quotidien, l'ensemble des activités, des interactions, des actions, des comportements, voire des pensées des individus seraient mis en données de façon à constituer un fonds numérique important. Ce « mini big data » personnel (Gurrian, Smeaton et Doherty, 2014) n'est plus seulement constitué de photographies, mais également d'écrits, de conversations, de chiffres, de métadonnées, etc. Cet ensemble peut servir à se rappeler des moments autobiographiques évanescents ou oubliés ou à faire des recoupements de données⁶⁴. Par la suite, ce vaste jeu de données pourrait participer à la création d'un « double » numérique de soi, comme le promet Personal AI, cas exemplaire en la matière. Sur la base d'un tel portrait comportemental, l'entreprise promet à ses client·es de créer des versions artificielles d'eux-mêmes sous la forme de robots conversationnels ; ces IA personnelles promettent de se comporter de la même façon que les individus humains qu'elles s'emploient à simuler, de travailler à leur place et de continuer d'interagir dans le monde après leur mort. Il n'est plus simplement question pour ces IA de prétendre copier certains phénomènes mnésiques comme ceux que nous avons évoqués précédemment (*anamnésis* et *mnémé*), mais de faire croire qu'elles peuvent reproduire le fonctionnement de la mémoire autobiographique d'un individu particulier et ainsi sa personnalité en vue « d'augmenter » la personne, voire l'humain lui-même. Une partie de l'industrie émergente de la recreation artificielle d'individus (Öhman et Floridi, 2017 ; Caccamo, 2021) repose sur la *datafication* continue et en temps réel du quotidien, avec une participation relativement volontaire des individus. On voit ainsi une gradation entre ce que proposent des technologies comme Google Photos et Apple Photos et un projet tel que celui de Personal AI où toute chose pourrait potentiellement être *datafiée*. Ces promesses répondent alors au « problème » de l'oubli — l'industrie est largement nourrie par un discours de refus de l'oubli : « Your AI Never Forgets⁶⁵ » (« Votre IA n'oublie jamais ») —, de la perte de la mémoire en étroit lien à la disparition des personnes. Passer par la promotion discursive de la préservation de la mémoire permet une plus grande acceptabilité sociale à l'égard de ce type de technologies que l'on peut qualifier de transhumanistes. La mémoire, ou plutôt sa production et sa conservation, sert ici d'argument fondamental pour légitimer une *datafication* en continu du quotidien. Ces promesses s'adosent également à une promesse épistémique bien connue suivant laquelle les « données » ne seraient pas des construits, mais aussi un reflet de la vie telle qu'elle a été « capturée⁶⁶ ». C'est en ce sens que Personal AI met en avant « un score d'authenticité » des IA qui tentent de copier les comportements des individus-sources : cet argument présent dans le discours de l'interface vise à convaincre de la légitimité de ces « doubles » numériques, à les rendre plus « vrais ».

Tableau 1. Vue d'ensemble de la rhétorique du problème et des promesses dans les discours des entreprises étudiées

	Google Photos	Apple Photos	Memoro	Personal AI
Rhétorique du « problème »				
Rhétorique du « problème » de trop-plein de données	x	x	x	
Rhétorique du « problème » de la faillibilité de la mémoire biologique (mémoire oublieuse)	x		x	x
Rhétorique du problème anthropologique (perte de la mémoire, effacement de la mémoire des personnes)				x
Promesses technomnésiques				
Augmenter la mémoire : <i>anamnèsis</i> algorithmisée (recherche active dans la mémoire avec augmentation du stockage)	x	x	x	x
Augmenter la mémoire : <i>mnémé</i> algorithmisée (évoquant spontanée)	x	x	x	x
— Rappel de mémoire épisodique (autobiographique)	x	x	x	
— Rappel de mémoire sémantique			x	

Conclusion

- 58 À travers l'examen critique des promesses technomnésiques produites par ces discours, cet article a donc rendu compte des processus de *datafication* et de la modélisation algorithmique de la mémoire sur lesquels reposent les promesses mnésiques formulées par Google Photos, Apple Photos, Memoro et Personal AI. La grande promesse d'automatisation de la mémoire est étroitement liée à une promesse d'augmentation de la mémoire. Cette dernière fait l'objet d'une variété de propositions en gradation : production, indexation et curation rationalisée de mémoire numérique (essentiellement visuelle et audiovisuelle), hiérarchisation du mémorable à la place et au bénéfice des individus, réponse à des prétendus besoins mnésiques et, en fin de compte, promesse d'une mémoire « quasi totale » qui pourrait mener à des versions de soi artificielles (IA personnelle) et à des « formes de vie » numériques *postmortem*.
- 59 Cet article a défendu l'idée que la mémoire algorithmique et ses récits génératifs ne constituent pas seulement une forme de *datafication* parmi d'autres, mais recouvrent une mise en données particulière qui transforme potentiellement le rapport qu'entretiennent les utilisateur·rice·s à leurs expériences du temps présent et de leurs souvenirs. Semblable à la photographie, qui fut au départ l'un des moyens de garder trace du présent pour devenir souvenir, le cadrage de nos expériences du temps présent devient celui d'une sauvegarde de données à l'infini dont la substance mnémonique sera prise en charge par les applications choisies. Il sera ainsi nécessaire de voir comment les personnes s'en servent concrètement et si les promesses sont vécues comme annoncées.
- 60 Au terme de cette enquête sur les promesses technomnésiques repérées dans les discours promotionnels de quatre entreprises proposant de prendre en charge, par différents moyens, la mémoire de leurs utilisateur·rice·s, il apparaît clairement, d'une part, que les discours qui soutiennent ces promesses participent au projet de production d'un « individu-data » projeté « dans un monde sans oubli » (Merzeau, 2013,

p. 115). Mais il apparaît également, d'autre part, que ces promesses produisent une certaine perte d'agentivité des utilisateur·ice·s dans la construction de leurs récits de vie. En effet, ces différentes technologies mnésiques reposent sur l'espoir d'une *datafication* quasi totale de l'expérience humaine et d'une interprétation efficace de ces données par des algorithmes présentés comme étant de plus en plus puissants, notamment par le biais du développement du *machine learning*. Cette organisation machinique des galeries et des contenus dignes d'être remémorés repose donc sur un ensemble de calculs probabilistes qui transforment la vie en données et la mémoire en matière statistique, écartant ce qui dans l'expérience humaine échappe encore à la *datafication* (les odeurs, le goût, le toucher et plus largement la mémoire procédurale sont exclus pour l'heure de cette mémoire algorithmique toute visuelle ou audiovisuelle) et aux critères parfois retors qui président à la survivance comme à la survenue de certains souvenirs (une association d'idées, le goût d'une madeleine trempée dans une tasse d'infusion). Alors que la métaphore de la boîte noire est très souvent mobilisée pour décrire l'opacité du fonctionnement des algorithmes, on peut opposer ici à leur logique statistique somme toute banale une autre boîte noire, autrement mystérieuse, qui est celle de la mémoire et de l'oubli humains.

BIBLIOGRAPHIE

- BARRETTE Pierre (2024), « La série Blue Moon, un laboratoire pour les séries originales de télévision par contournement (TPC) locales », *Kinephanos*, 10(1), p. 70-93.
- BAZIN André (1981), « Ontologie de l'image photographique », dans *Qu'est-ce que le cinéma ?*, Paris, Les Éditions du Cerf, p. 9-17.
- BEER David et Roger BURROWS (2013), « Popular culture, digital archives and the new social life of data », *Theory, Culture & Society*, 30(4), p. 47-71, DOI : <https://doi.org/10.1177/0263276413476542>.
- BHOWMIK Samir (2019), « Thermocultures of memory », *Culture Machine*, 17, p. 1-20.
- BOUDON Héroïse et Katharina NIEMEYER (2016), « Fidelity as interseriality: The challenges of adapting the Québécois TV series *Les Invincibles* for broadcast in France », *Canadian Journal of Communication*, 41(4), p. 655-672.
- CACCAMO Emmanuelle (2021), « Eternime ou simuler les morts par le biais d'agents conversationnels "intelligents". Réflexions sémiotiques sur un média imaginaire », *Frontières*, 2(32), DOI : <https://doi.org/10.7202/1083220ar>.
- CACCAMO Emmanuelle (2022), Répertoire Technomnésie, <https://technomnesia.org/repertoire/>, page consultée le 9 octobre 2025.
- CACCAMO Emmanuelle, Océane GUERET et Phuong Anh NGUYEN (2026), « "J'ai trouvé une app' qui peut enregistrer votre vie automatiquement !" : Lumière sur les motivations, les croyances et les usages de *lifeloggers* », dans Emmanuelle CACCAMO et Katharina NIEMEYER (dir.), *Technomémoires*, Québec, Presses de l'Université du Québec, p. 149-174.

- CHENEY-LIPPOLD John (2017), *We are data: Algorithms and the making of our digital selves*, New York, NYU Press.
- CONNERTON Paul (2008), « Seven types of forgetting », *Memory Studies*, 1(1), p. 59-71.
- CRÉMIER Lucile, Maude BONENFANT et Laura I. LAFRANCE ST-MARTIN (2019), « Données brutes ou hypersymboles ? Signification et données numériques, entre processus discursif et procédure machinique », *Cygne noir*, 7, p. 79-110, DOI : <https://doi.org/10.7202/1089330ar>.
- DANDURAND Guillaume, Florence LUSSIER-LEJEUNE, Daniel LETENDRE et Marie-Jean MEURS (dir.) (2022), *Attentes et promesses technoscientifiques*, Montréal, Presses de l'Université de Montréal.
- DAVALLON Jean (1984), « Sociosémiotique des images », *Langage & société*, 28(2), p. 111-140.
- DRAAISMA Douwe (2000), *Metaphors of memory*, Cambridge, Cambridge University Press.
- DENIS Jérôme (2018), *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines.
- DODGE Martin et Rob KITCHIN (2007), « Outlines of a world coming into existence: Pervasive computing and the ethics of forgetting », *Environment and Planning B Planning and Design*, 34(3), p. 431-445.
- ERLL Astrid (2011), « The invention of cultural memory: A short history of memory studies », dans *Memory in culture*, Londres, Palgrave Macmillan, p. 13-37.
- FAN Chenyou, Zehua ZHANG et David J. CRANDALL (2008), « Deepdiary: Lifelogging image captioning and summarization », *Journal of Visual Communication and Image Representation*, 55, p. 40-55.
- GARDE-HANSEN Joanne, Andrew HOSKINS et Anna READING (dir.) (2009), *Save as... digital memories*, New York, Palgrave Macmillan.
- GARDE-HANSEN Johanne (2011), *Media and memory*, Édimbourg, Edinburgh University Press.
- GURRIN Cathal, Alan F. SMEATON et Aiden R. DOHERTY (2014), *Lifelogging: Personal big data*, Boston, Now.
- HOGAN Mél (2024), « The fumes of AI », *Critical AI*, (2)1, DOI : <https://doi.org/10.1215/2834703X-11205231>.
- HOSKINS Andrew (2024), « AI and memory », *Memory, Mind & Media*, 3, DOI : <https://doi.org/10.1017/mem.2024.16>.
- JACOBSEN Ben et David BEER (2021), *Social media and the automatic production of memory: Classification, ranking and the sorting of the past*, Bristol, Bristol University Press.
- JACOBSEN Benjamin N. (2020), « “You Have a New Memory”: Mediated memories in the age of algorithms », Toronto, University of York, https://etheses.whiterose.ac.uk/28341/1/Jacobsen_109033505_Thesis.pdf, page consultée le 10 octobre 2025.
- JOLY Pierre-Benoît (2013), « À propos de l'Économie des promesses techno-scientifiques », dans Jacques LESOURNE et Denis RANDET (dir.), *La recherche et l'innovation en France FuturIS*, Paris, Odile Jacob, p. 231-255.
- KEUCHEYAN Razmig (2019), *Les besoins artificiels : comment sortir du consumérisme*, Paris, Zones.
- KIDD Jenny et Eva NIETO MCAVOY (2023), « Deep nostalgia: Remediated memory, algorithmic nostalgia and technological ambivalence », *Convergence*, 29(3), p. 620-640.

- LEE JeongHyun (2020), « Algorithmic uses of cybernetic memory: Google Photos and a genealogy of algorithmically generated “memory” », *Social Media + Society*, 6(4), DOI : 10.1177/2056305120978968.
- LIV Nadine et Dov GREENBAUM (2020), « Deep fakes and memory malleability: False memories in the service of fake news », *AJOB Neuroscience*, 11(2), p. 96-104.
- MERZEAU Louise (2013), « L’intelligence des traces », *Intellectica. Revue de l’Association pour la recherche cognitive*, 59(1), p. 115-135, DOI : <https://doi.org/10.3406/intel.2013.1088>.
- MILLETTE Mélanie, David MYLES, Florence MILLERAND et al. (2020), *Méthodes de recherche en contexte numérique : une orientation qualitative*, Montréal, Presses de l’Université de Montréal.
- MORELLI Pierre et Mirela LAZAR (2015), « Plasticité des dispositifs d’information et de communication », *Questions de communication*, 28, p. 7-17, DOI : <https://doi.org/10.4000/questionsdecommunication.10017>.
- NIEMEYER Katharina (2011), *De la chute du mur de Berlin au 11 Septembre 2001. Le journal télévisé, les mémoires collectives et l’écriture de l’histoire*, Lausanne, Antipodes.
- NIEMEYER Katharina et Chloé DOUAUD (2018), « La marchandisation du passé : le cas d’Apple et de la préfabrication des souvenirs », *Effeuillage*, 7, p. 78-81, DOI : <https://doi.org/10.3917/eff.007.0078>.
- NIEMEYER Katharina et Valérie SCHAFER (2022), « Stocker. Penser et placer le stockage au cœur des pratiques médiatiques et communicationnelles », *Le Temps des médias*, 39(2), p. 6-14, DOI : <https://doi.org/10.3917/tdm.039.0006>.
- ÖHMAN Carl et Luciano FLORIDI (2017), « The political economy of death in the age of information: A critical approach to the Digital Afterlife Industry », *Minds and Machines*, 27(4), p. 639-662.
- PAILHES Amandine et Giuseppina SAPIO (2022), « Une photo à soi et des errances sémiologiques », *Communication & langages*, 2(212), p. 63-75.
- POITOU Jean-Pierre (1999), « Histoire de la mémoire artificielle », dans Charles LENAY et Véronique HAVELANGE (dir.), *Mémoire de la technique et techniques de la mémoire*, Paris, Érès, p. 35-61.
- RICEUR Paul (2003), *La mémoire, l’histoire, l’oubli*, Paris, Seuil.
- ROUGE Rémi (2019), « Des souvenirs très précieux : souvenirs numériques et valorisation des données », dans Éric GEORGE (dir.), *Numérisation de la société et enjeux sociopolitiques 1*, Londres, ISTE Éditions, p. 89-98.
- ROUGE Rémi (2023), *Le temps retrouvé, la fabrique des souvenirs numériques saisie par ses professionnels*, thèse de doctorat, Paris, Université Paris 8, <https://theses.fr/2023PA080031>, page consultée le 9 octobre 2025.
- ROUGE Rémi (à paraître), « Interpréter les données pour produire des souvenirs numériques : l’exemple de la sélection affective », *Études digitales*, 17.
- SCHINCARIOL Alexa, Henry OTGAAR, Ciara M. GREENE et al. (2024), « Fake memories: A meta-analysis on the effect of fake news on the creation of false memories and false beliefs », *Memory, Mind & Media*, 3(e17), DOI : 10.1017/mem.2024.14.
- SELKE Stefan (dir.) (2016), *Digital self-tracking and lifelogging – Between disruptive technology and cultural transformation*, Wiesbaden, Springer.

SMIT Rik (2024), « When memories become data », dans *The remaking of memory in the age of the Internet and social media*, Oxford, Oxford University Press, p. 196-209.

STIEGLER Bernard (1991), « État de la mémoire et mémoire de l'État », *Ars industrialis*, <https://arsindustrialis.org/node/1936>, page consultée le 9 octobre 2025.

TODOROV Tzvetan (2004), *Les abus de la mémoire*, Paris, Arléa.

TRIGG Dylan (2009), « The place of trauma: Memory, hauntings, and the temporality of ruins », *Memory Studies*, 2(1), p. 87-101.

VAN DIJCK José (2014), « Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology », *Surveillance & Society*, 12(2), p. 197-208, DOI : <https://doi.org/10.24908/ss.v12i2.4776>.

VILVANG Chrys (2024), « Between automated memory and history: Blocking “sensitive locations” from Apple Memories », *Memory, Mind & Media*, 3(8), DOI : <https://doi.org/10.1017/mem.2024.4>.

WANG Qi et Andrew HOSKINS (2024), *The remaking of memory in the age of the Internet and social media*, Oxford, Oxford University Press.

WEST Nancy Martha (2000), *Kodak and the lens of nostalgia*, Charlottesville, University of Virginia Press.

ZULFIKAR Wazeer, Samantha CHAN, Pattie MAES (2024), « Memoro: Using large language models to realize a concise interface for real-time memory augmentation », *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, DOI : <https://doi.org/10.1145/3613904.3642450>.

NOTES

1. Google Photos, « Your Photos, your memories, your way », Blog Google – Photos (blog), 18 mai 2021, <https://blog.google/products/photos/new-memories-features-look-back/>.
2. Limitless : <https://www.limitless.ai/> ; Omi : <https://www.omi.me/> ; Plaud : <https://ca.plaud.ai/> ; Memoro : <https://www.media.mit.edu/projects/memoro/overview/>.
3. Le terme *données* peut renvoyer à un ensemble d'éléments très variés – textes, photographies, captures d'écran, vidéos, dessins, enregistrements vocaux et audio, données corporelles, données de localisation et d'horodatage de la création des contenus, etc. – à partir desquels les algorithmes opèrent des calculs (classification, détection de patterns, etc.). Le traitement algorithmique de nos archives numériques par les technologies mnésiques suppose donc qu'elles soient préalablement *datafiées*. Comme l'explique Cheney-Lippold (2017, p. 13), un ordinateur ne reconnaît pas une personne joyeuse sur une photo, il encode l'image en procédant à des calculs qui produisent, in fine, un « sourire-“données” » (« data-“smile” », traduction libre).
4. Les *memory studies* s'intéressent également à la mémoire traumatique (Trigg, 2009), mais aussi, notamment dans le contexte de la *datafication*, aux enjeux écologiques du stockage des données (Hogan, 2024), au potentiel du devenir historique des données du présent (Niemeyer et Schafer, 2022), aux phénomènes de réinvention accrue du passé liés aux fausses nouvelles (Schincariol et al, 2024), aux agents conversationnels automatisés (Hoskins, 2024) ainsi qu'à la place de la mémoire dans un environnement surconnecté (Wang et Hoskins, 2024).
5. Pierre-Benoît Joly soutient que les promesses scientifiques ont « une dimension discursive (mots d'ordre, arguments) et une dimension matérielle (pratiques, organisations sociales, objets physiques...) » et qu'elles « interagissent » (*op. cit.*, p. 241). Cela est d'autant plus vrai que la dimension discursive est elle-même le produit de pratiques professionnelles.

6. Il convient de souligner que chaque innovation technologique dans le domaine de la mémoire véhicule un ensemble de promesses spécifiques. L'avènement du phonographe, de la photographie et du cinématographe s'est par exemple accompagné de la promesse d'une restitution de la réalité dans son immédiateté quasi brute (voir notamment Bazin, 1981). L'ouvrage *Technomémoires* édité par Emmanuelle Caccamo et Katharina Niemeyer (sous presse) prolongera ces réflexions.

7. Le terme *mnésique* (ou mnémonique), du grec *mnêsios* ou du latin *mens*, signifie « de mémoire » et renverrait également au mot grec *mnasthai* (se souvenir). Voir <https://www.cnrtl.fr/definition/academie9/mn%C3%A9sique>, dernier accès le 18 février 2025.

8. Depuis 2015, Google/Alphabet.

9. Ces technologies correspondent à des formes assez différentes : assistants mnésiques (Memoro, Personal AI), albums photo « intelligents » (Apple Photos, Google Photos) et fonctionnalités mnésiques (Apple Souvenir, Google Souvenir). Ce choix correspond à notre souhait de couvrir un ensemble de promesses. Pour des raisons évidentes d'espace, nous nous sommes limitées à ces cas parmi le grand nombre d'autres technologies disponibles (voir Caccamo, 2022 ; Rouge, 2023).

10. <https://blog.google/>.

11. <https://www.media.mit.edu/projects/memoro/overview/>.

12. <https://www.personal.ai/thesis>.

13. <https://www.personal.ai/memory>.

14. <https://docs.personal.ai/docs/terminology> ; <https://www.personal.ai/train>.

15. Il y a différentes manières de définir la sociosémiotique. Dans cet article, cette dernière est comprise et approchée dans sa dimension critique et ouverte en permettant de décoder les signes visuels, textuels et procéduraux des textes, des sites Web et des interfaces tout en leur accordant la force de coconstruire l'univers social, économique et culturel de la production et de la réception dans sa discursivité et visualité ; univers duquel les signes sont eux aussi une construction (Barrette, 2024). Certaines chercheuses ne mobilisent pas d'entrevues dans leurs recherches sociosémiotiques en défendant l'idée que les textes et images analysés, coproduisant le social, tout en le constituant, ne nécessitent pas d'autres méthodologies. Bien qu'en raison de l'angle choisi dans cet article nous ne mesurons pas ici l'écart entre les promesses formulées et leur efficacité effective, l'ouverture méthodologique est bénéfique pour renforcer certains résultats de l'analyse textuelle et visuelle ou encore pour démontrer que le sens n'est pas toujours celui qui semble se construire au prisme de la seule lecture des chercheurs (Boudon et Niemeyer, 2016).

16. Ce constat fait l'objet d'un argument commercial dans le discours de Personal AI qui se positionne comme une nouvelle forme de technologie de la mémoire : <https://www.personal.ai/thesis>.

17. « Répondre à un besoin » est une reformulation de la prétention à « résoudre un problème ». Les deux formulations permettent de présenter l'activité des professionnel·les comme une activité visant à « aider » les utilisateur·rice·s en leur fournissant un « service ». On trouve régulièrement, dans l'explication que fournissent les concepteur·rice·s de ces technologies en entretien (Rouge, 2023), l'argument selon lequel elles émergent comme solution au problème de la gestion de contenus numériques dont le volume va croissant.

18. Joly note qu'en effet, la production d'une promesse technoscientifique s'articule toujours à la production d'une légitimité : « Parmi les ressources utilisées pour intéresser un large public, on trouve ainsi des outils rhétoriques généraux. Car la promesse n'a pas seulement à être crédible, elle doit aussi être légitime » (*op. cit.*, p. 237).

19. <https://blog.google/products/photos/>.

20. Traduction libre : « Plus nous capturons de moments, plus il devient difficile de revivre ces souvenirs. Les photos et vidéos s'entassent sur nos appareils mobiles, nos anciens ordinateurs, nos disques durs et nos services en ligne (qui manquent constamment d'espace). [...] Nous

voulions faire mieux. Alors aujourd’hui, nous présentons Google Photos — un nouveau produit autonome qui vous offre un foyer pour toutes vos photos et vidéos, qui vous aide à organiser et à donner vie à vos moments, et vous permet de partager et sauvegarder ce qui compte », <https://blog.google/products/photos/picture-this-fresh-approach-to-photos/>.

21. <https://blog.google/products/photos/redesigned-google-photos/>.

22. Traduction libre : « Souvent, nous prenons des photos pour capturer des souvenirs ou sauvegarder des détails importants, mais les informations contenues dans ces photos peuvent se mélanger. [...] Les capacités multimodales de Gemini peuvent comprendre le contexte et le sujet des photos pour en extraire les détails », <https://blog.google/products/photos/ask-photos-google-io-2024/>. On trouve également la justification de l’usage de l’IA comme solution au trop-plein de contenus dans l’article de 2021 cité en accroche de cet article : <https://blog.google/products/photos/new-memories-features-look-back/>.

23. <https://www.apple.com/ca/macros/photos/> et <https://www.apple.com/ca/fr/visionos/visionos-2/>.

24. Ce récit s’appuie donc sur la nature pharmacologique de la technologie, au sens où l’entend Stiegler, <https://arsindustrialis.org/vocabulaire-pharmakon-pharmacologie>.

25. Traduction libre : « Alors que vous continuez à prendre des photos et à capturer des vidéos, nous continuerons à travailler pour les rendre encore plus faciles à stocker, à partager et à faire revivre », <https://blog.google/products/photos/picture-this-fresh-approach-to-photos/>.

26. Traduction libre : « Au fond, la rétention de la mémoire est un problème biologique. Mais au lieu de se concentrer sur la réparation des personnes avec des médicaments d’amélioration cérébrale ou des implants chirurgicaux, pourquoi ne pas construire un outil pour résoudre le problème ? » Personal AI ajoute également : « You don’t solve the human, you solve the problem » (« On ne répare pas l’humain, on résout le problème »), <https://www.personal.ai/thesis>.

27. Les discours transhumanistes ont participé à diffuser l’idée d’une mémoire originellement handicapée (et du corps humain au complet), comme un problème à résoudre par des technologies émergentes. Dans ce type de représentations sociales, la perte de mémoire est également associée aux thèmes de la perte d’identité et de la mort, de la mort du sujet ou de « micro-morts » de soi. Sur ce dernier thème, voir Caccamo, Gueret et Nguyen (à paraître).

28. Traduction libre : « Depuis le début de l’histoire humaine, la mémoire d’une personne pouvait être perdue ; maintenant, à l’ère numérique — et plus spécifiquement, à l’ère de l’IA —, votre mémoire peut être rappelée et partagée de façon fluide », <https://www.personal.ai/memory>.

29. « Real-Time Memory Augmentation » (Zulfikar, Chan et Maes, 2024).

30. Traduction libre : « Les individus doivent mémoriser un volume d’informations en constante expansion. Pour les aider, nous avons développé Memoro, un assistant personnel de mémoire portable et basé sur l’audio qui utilise un agent d’IA. Memoro capture automatiquement les interactions auditives du monde réel de l’utilisateur et les convertit en souvenirs numériques. Lors du rappel, l’assistant peut directement déduire les besoins de mémoire en utilisant le contexte, effectuer des recherches et présenter des suggestions minimales à intégrer dans la vie quotidienne et les conversations », <https://www.media.mit.edu/projects/memoro/overview/>, dernière consultation le 15 décembre 2024.

31. On notera également dans cette citation la mention du problème d’un trop grand volume d’informations et la rhétorique du besoin.

32. <https://blog.google/products/photos/give-and-get-photos-you-care-about/>.

33. Voir, par exemple, Poitou (1999).

34. Traduction libre : « Vos souvenirs automatiquement organisés ».

35. <https://blog.google/products/photos/gallery-go/>.

36. Traduction libre : « Un album photo plus intelligent, sans effort ».

37. Traduction libre : « Google Photos a été l'un des premiers produits que nous avons conçus en mettant l'IA au centre, vous donnant la possibilité de rechercher dans vos photos et vidéos des personnes, des animaux de compagnie, des lieux et bien plus encore », <https://blog.google/products/photos/ask-photos-google-io-2024/>.
38. Souvent appelées données EXIF (pour « *exchangeable image file format* »), elles peuvent comprendre le lieu, la date de la prise de vue, le type d'appareil utilisé, voire ses caractéristiques et réglages.
39. <https://blog.google/products/photos/smarter-photo-albums-without-work/>.
40. Dans un article consacré aux racines cybernétiques de la mémoire algorithmiquement produite, à l'exemple de Google Photos, JeongHyun Lee (2020, p. 6) revient sur l'usage des métadonnées pour l'organisation chronologique des albums et sur le développement d'algorithmes de *machine learning* visant à identifier des « objets » sur des photos générant à leur tour des métadonnées permettant « de catégoriser et d'organiser toutes les photos pour qu'elles puissent être retrouvées dans le futur » (p. 7). Traduction libre de « *categorizing and organizing all photos for future retrievals* ».
41. <https://blog.google/products/photos/google-ai-photo-editing-features-tips/>, <https://blog.google/products/photos/google-ask-photos-early-access/>.
42. Voir, par exemple, le classement des « thèmes » de mémoire effectué par Facebook tels qu'ils sont relatés dans <https://research.facebook.com/blog/2017/9/facebook-memories-the-research-behind-the-products-that-connect-you-with-your-past/>.
43. Traductions libres : « Vous perdez cette sensation chaleureuse et nostalgique lorsque vous devez faire défiler des centaines de photos en double, alors nous mettons vos souvenirs au premier plan dans Google Photos » ; « Pour faciliter le retour en arrière, nous utilisons l'IA pour alimenter de nouvelles fonctionnalités qui font ressurgir les moments significatifs et donnent vie à vos souvenirs [...] » <https://blog.google/products/photos/new-memories-features-look-back/>.
44. Traduction libre : « Nos téléphones sont remplis d'un nombre croissant de photos et de vidéos capturant tout, des moments du quotidien aux étapes importantes de la vie. Mais ils sont aussi encombrés d'images en double, de captures d'écran et de vidéos floues, ce qui rend difficile de se concentrer sur les moments les plus significatifs », <https://blog.google/products/photos/google-photos-memories-view/>.
45. Traduction libre : « Google Photos est le foyer de vos souvenirs. Mais nous savons que votre galerie photo peut s'encombrer de photos similaires d'un même moment ou de captures d'écran et de documents que vous avez pris en photo pour référence future », <https://blog.google/products/photos/google-photos-organization-updates-november-2023/>.
46. <https://blog.google/products/photos/how-google-photos-best-take-works/>.
47. Traduction libre : « Nous utilisons l'apprentissage automatique pour sélectionner ce qui apparaît dans Souvenirs, afin que vous n'ayez pas à parcourir de nombreuses prises identiques et que vous puissiez plutôt vous concentrer sur les meilleures, où les photos sont de bonne qualité et où tous les enfants sourient », <https://blog.google/products/photos/relive-your-best-memories-new-features-google-photos/>.
48. Traduction libre : « Les fêtes sont toujours un bon moment pour réfléchir à ce qui compte vraiment — pour moi, c'est ma famille, celle qui est proche et celle restée en Europe. Avec Souvenirs dans Google Photos, je peux revivre les moments — grands et petits — que j'ai vécus avec mes proches », <https://blog.google/products/photos/new-cinematic-photos-and-more-ways-relive-your-memories/>.
49. <https://blog.google/products/photos/new-memories-features-look-back/>.
50. Traduction libre : « Avec Souvenirs, vous pouvez déjà redécouvrir les photos importantes des années passées, les moments forts récents, les instants avec vos proches, vos activités préférées et bien plus encore. Grâce à l'apprentissage machine, nous pouvons maintenant proposer plus que faire simplement ressortir des photos basées sur des thèmes pour le faire en nous appuyant

sur des motifs visuels particuliers dans vos photos. À partir de cet été, lorsque nous trouvons un ensemble de trois photos ou plus qui partagent des éléments comme la forme ou la couleur, nous mettrons en évidence ces petits motifs pour vous dans vos Souvenirs. Par exemple, l'un de nos ingénieurs a reçu cette collection présentant les photos qu'il avait prises de son sac à dos orange préféré », <https://blog.google/products/photos/new-memories-features-look-back/>.

51. « Important » et « meaningful », <https://blog.google/products/photos/new-memories-features-look-back/>. Chez Apple, le terme *significatif* semble appliqué lors d'un déplacement dans un lieu inédit, lorsque plusieurs images ont été prises au domicile durant une période de temps (« Novembre à la maison ») ou à la prise de photos lors d'un événement calendaire : une photo prise dans un nouveau lieu, quelques heures passées à un nouvel endroit ou un album d'images constitué à une date culturellement importante (Noël par exemple) peuvent suffire à être catégorisés comme « significatifs ». L'acception demeure assez réductionniste.

52. « Moments that matter », <https://blog.google/products/photos/turn-frown-upside-down-suggested-rotations-and-more/>.

53. Traductions libres : « meilleurs moments », « meilleurs souvenirs », <https://blog.google/products/photos/relive-your-best-memories-new-features-google-photos/>.

54. Traduction libre : « Vos personnes, lieux et activités préférés », <https://blog.google/products/photos/new-cinematic-photos-and-more-ways-relive-your-memories/>.

55. Les utilisateur·rice·s peuvent faire une expérience appauvrie de leur mémoire, cadrée par un ensemble de catégories prédéfinies par un tiers. Dans le vaste fonds de photographies prises par les individus, seules certaines catégories sont considérées comme « valables ». Ces choix participent à un oubli structurel (Connerton, 2008), au sens où, parmi la masse de contenus accumulés, certains peuvent tomber dans l'oubli, car ils ne sont pas ramenés « à la surface » de l'interface, bien que présents, retrouvables.

56. Traduction libre : « Prises individuellement, ces photos ne sont peut-être pas les plus significatives, mais quand vous les voyez toutes ensemble, elles racontent une histoire — votre histoire », <https://blog.google/products/photos/new-memories-features-look-back/>.

57. « Fairness in the Generation of Face Embeddings. A challenge in obtaining a useful face representation is ensuring consistent accuracy on many axes. The model must show similar performance across various age groups, genders, ethnicity, skin tones and other attributes. Fairness is an essential aspect of model development and must be taken into account from the beginning, not only in the data collection which needs to be diverse, inclusive, and balanced but also in strong failure analysis and model evaluation. » Traduction libre : « Équité dans la génération d'*embeddings* faciaux. Pour obtenir une représentation faciale utile, le défi consiste à garantir une précision cohérente selon plusieurs axes. Le modèle doit montrer des performances similaires à travers différents groupes d'âge, genres, ethnies, teintes de peau et autres attributs. L'équité est un aspect essentiel du développement de modèles et doit être prise en compte dès le début, non seulement dans la collecte de données qui doit être diversifiée, inclusive et équilibrée, mais aussi dans une analyse rigoureuse des défaillances et l'évaluation du modèle », <https://machinelearning.apple.com/research/recognizing-people-photos>.

58. <https://machinelearning.apple.com/research/recognizing-people-photos>.

59. Traduction libre : « Les photos cinématiques vous aident à revivre vos souvenirs d'une manière qui semble plus vivante et réaliste — pour que vous ayez l'impression d'être transporté dans ce moment passé », <https://support.google.com/photos/thread/88778155/relive-the-moment-with-cinematic-photos?hl=en>.

60. <https://www.apple.com/ca/fr/visionos/visionos-2/>.

61. <https://www.apple.com/ca/fr/visionos/visionos-2/>.

62. Cela peut faire penser au projet annoncé en 2018 par Facebook de créer des environnements mnésiques en réalité virtuelle et en 3D à partir de photographies : https://www.youtube.com/watch?v=Ur1Z72_LyTM&ab_channel=PetaPixel.

63. Pour une analyse plus détaillée de cette publicité, voir Niemeyer et Douaud (*op. cit.*) et Smit (*op. cit.*).
64. Sur les usages du *lifelogging*, voir Caccamo, Gueret et Nguyen, 2026.
65. <https://www.personal.ai/train>, dernière consultation le 17 avril 2024.
66. Sur les modes de représentation des données (fausse épistémie), voir Crémier, Bonenfant, Lafrance St-Martin (2019). La diffusion de la photographie a elle aussi, en son temps, été accompagnée d'une croyance épistémique ou ontologique forte. Voir par exemple Bazin (*op. cit.*) sur les croyances associées à l'ontologie de l'image photographique.

RÉSUMÉS

Cet article examine les promesses technomnésiques formulées dans les discours d'entreprises développant des technologies de *datafication* de la mémoire personnelle (Google Photos, Apple Photos, Memoro, Personal AI). À travers une analyse des discours promotionnels de ces entreprises, l'article identifie une promesse générale d'automatisation des processus mnésiques, présentée comme solution à deux « problèmes » que ces entreprises ont contribué à construire : le trop-plein de données numériques et la faillibilité de la mémoire humaine. Cette promesse générale se décline en sous-promesses : l'« augmentation » de la mémoire par l'automatisation de la recherche active (*anamnèsis* algorithmisée) et de l'évocation spontanée de souvenirs (*mnémé* algorithmisée) ainsi que la création d'un « double » numérique de soi par la *datafication* continue du quotidien. L'article montre que ces technologies opèrent une sélection normative du mémorable selon des critères qui ne répondent pas tant aux besoins des utilisateur·rice·s qu'à ceux (techniques, commerciaux, etc.) des entreprises qui les produisent. La mémoire personnelle se trouve ainsi transformée en matière statistique, excluant les dimensions sensorielles et les processus mnésiques complexes qui échappent encore à la *datafication*.

This article examines the memory datafication promises made in the discourses of companies developing technologies that datafy personal memory (Google Photos, Apple Photos, Memoro, Personal AI). By analyzing the promotional discourses of these companies, the article identifies a general promise of automating memory processes, presented as a solution to two “problems” that these companies have helped create: the overload of digital data and the fallibility of human memory. This general promise is broken down into subpromises: the “augmentation” of memory through the automation of active searching (algorithmic *anamnesis*) and spontaneous recall of memories (algorithmic *mnemonics*), as well as the creation of a digital “double” of the self through the continuous datafication of daily life. The article shows that these technologies perform a normative selection of what is memorable according to criteria that cater not so much to the needs of users, but to the technical, commercial, and other interests of the companies that produce them. Personal memory is thus transformed into statistical data, excluding the sensory dimensions and complex mnemonic processes that remain beyond the reach of datafication.

Este artículo presenta un estudio de las promesas tecnomnésicas expresadas en los discursos de las empresas que desarrollan tecnologías de *dataficación* de la memoria personal (Google Fotos, Apple Fotos, Memoro, Personal IA). Por intermedio de un análisis de los discursos de promoción de estas empresas, el artículo identifica un compromiso general de automatización de los procesos de nemotecnia, presentados como solución a dos “problemas” que esas empresas han

generado, es decir, el exceso de los datos digitales y la falibilidad de la memoria humana. Este compromiso general se declina en subcompromisos, es decir, el “aumento” de la memoria por la automatización de la investigación activa la anamnesis algorítmica y la evocación espontánea de los recuerdos (*nemónico* algorítmico), así como la creación de un “doble” digital de sí mismo por intermedio de la *dataficación* continua en el cotidiano. El artículo demuestra que esas tecnologías realizan una selección normativa de lo inolvidable, según criterios que no responden mucho a las necesidades de los usuarios que a aquellos (técnicos, comerciales, etc.), de las empresas que las producen. La memoria personal se encuentra esa manera, transformada en materia estadística, excluyendo las dimensiones sensoriales y los procesos mnemotécnicos complejos que todavía escapan a la *dataficación*.

INDEX

Mots-clés : mémoire personnelle, souvenirs, automatisation, promesses, discours promotionnels

Keywords : automation, promotional discourses, personal memory, promises, memories

Palabras claves : automatización, discursos promocionales, memoria personal, promesas, recuerdos

AUTEURS

RÉMI ROUGE

Rémi Rouge est docteur, enseignant contractuel en sciences de l'information et de la communication à l'Université Sorbonne Paris Nord, CRESSPA. Courriel : rougeremi@gmail.com

EMMANUELLE CACCAMO

Emmanuelle Caccamo est professeure à l'Université du Québec à Trois-Rivières. Courriel : emmanuelle.caccamo@uqtr.ca

KATHARINA NIEMEYER

Katharina Niemeyer est professeure à l'Université du Québec à Montréal. Courriel : niemeyer.katharina@uqam.ca

La datafication des soins en maisons de santé

Un pouvoir professionnel inégalement distribué

Manon Plegat

- 1 Situées à l'interface entre l'intimité du corps (Dagiral *et al.*, 2019 ; Lupton, 2017) et la dimension publique de la santé (Chambaud et Schwartz, 2016), les données de soins occupent une position spécifique dans les représentations sur le numérique. Leur utilisation par des acteurs privés, tels que les assurances, ou publics, comme les États, suscitent des craintes formulées en termes de marchandisation critiquable (Jeanningros, 2021 ; McFall, 2019) ou de gouvernement des corps questionnable (Fassin et Memmi, 2004). En contraste, le « mandat » (Hughes, 1997) dont disposent les soignants en ferait des acteurs plus légitimes pour la production de données médicales. Leur engagement formel à l'égard de la santé des patients et leur expertise professionnelle en garantiraient un usage parcimonieux et pertinent¹. Cependant, en France, malgré ces prédispositions, l'exploitation de bases de données par les praticiens eux-mêmes fait face à de nombreux freins.
- 2 Dans le secteur hospitalier, où les méthodes du nouveau management public sont mises en place depuis plus de deux décennies, les informations tracées sont largement associées à des objectifs gestionnaires (Bertillot, 2014) en tension avec la production de savoirs cliniques (Juven, 2018). Dans le secteur libéral, la même logique de rationalisation se déploie (Da Silva, 2017), mais avec une force moindre. Si, en médecine de ville, les soignants doivent attester de leur bon respect des recommandations afin d'obtenir des rémunérations dédiées, le contrôle exercé par les autorités est pour partie décorrélé de leurs pratiques d'encodage des dossiers patients (Bloy et Rigal, 2012). Nombre des indicateurs d'évaluation sont soit produits par les services de sécurité sociale eux-mêmes (à partir des informations de facturation télétransmises à l'assurance maladie), soit déclarés par les professionnels. Ils ne nécessitent donc pas systématiquement la production, par les praticiens, de données exploitables informatiquement. Ces quelques éléments soulignent l'existence d'un cloisonnement entre les pratiques du soin et les conditions nécessaires à l'analyse scientifique.

- 3 Cependant, depuis 2007, le développement des maisons de santé pluriprofessionnelles (MSP) semble avoir ouvert une voie alternative aux bases de données médico-administratives des hôpitaux et aux « non-bases de données » de la médecine de ville — où les informations ne sont pas structurées, car la saisie en texte libre est la plus courante. À rebours des clivages inhérents au système des professions (Abbott, 1988) de santé, ces organisations regroupent des soignants libéraux issus de métiers différents afin qu'ils coordonnent leur prise en charge des patients. Désormais réunis en équipes de soins primaires, les praticiens sont invités à utiliser un système d'information partagé (SIP) labellisé par les pouvoirs publics. Pour qu'un établissement soit légalement reconnu comme une MSP, les soignants doivent choisir leur logiciel parmi une liste de 16 SIP préidentifiés par l'Agence du Numérique en Santé² comme étant « pluriprofessionnels », et donc adaptés à ce nouveau mode d'organisation des soins.
- 4 Commun à tous les soignants, le système labellisé doit centraliser les données sur les patients, mais aussi accompagner leur encodage. En multipliant le nombre de praticiens rédacteurs et en incitant à la standardisation de leurs écritures, le logiciel partagé est censé permettre la constitution de bases informatiques exhaustives et originales. Selon ses promoteurs, il aiderait à « la structuration des informations médicales et [améliorerait] leur qualité » (ASIP Santé, 2016a). À la différence des données recueillies en laboratoire ou enregistrées par les organismes de sécurité sociale, l'outil labellisé favoriserait ainsi le développement de statistiques « locales » aptes à révéler les besoins « des territoires ». Les SIP labellisés sont donc associés à une promesse technoscientifique (Joly, 2010) propice à la conduite d'une étude de la *datafication*, où la mise en machine et la mise en chiffres se répondent l'une l'autre. La politique qui les promeut est doublement marquée par le registre du numérique : l'utilisation d'une même technologie numérique par tous les soignants permettrait la production de données, numériques elles aussi. Les SIP prescrits seraient un outil-support à la transcription quantifiée des activités des équipes pluriprofessionnelles ou, autrement dit, un outil propice à la *datafication* des soins (Mayer-Schönberger et Cukier, 2014). Une fois informatisées et agrégées, les informations sur la santé des patients seraient exploitables à des fins de recherche.
- 5 Toutefois, le raisonnement de la politique publique étudiée est en décalage avec les résultats produits par les analyses critiques de la mise en machine et de la mise en chiffres. Les liens entre l'outil, son environnement et ses usages ont été pensés comme un processus mécanique. Pourtant, de nombreux travaux en sciences sociales ont insisté sur les limites des approches technodéterministes, considérant que l'unicité de l'instrument entraînerait automatiquement l'unification des pratiques. Appelant à appréhender les « données » comme des « obtenues » (Latour, 2007, p. 188), ils ont aussi attiré l'attention sur le rôle crucial que jouent les travailleurs pour faire des langages, espaces et informations initialement éclatés des variables commensurables (Heaton et Proulx, 2012). De telles analyses ont par ailleurs montré que ce travail des données (Bossen *et al.*, 2019 ; Denis, 2018) tendait à être invisibilisé selon la position hiérarchique de l'utilisateur dans le processus de production numérique (Bossen *et al.*, 2014 ; Casilli, 2019/2021 ; Millerand, 2012). Ces enseignements sont à l'origine de nos réflexions, puisqu'ils font écho aux caractéristiques intrinsèques de notre objet de recherche. Les pratiques en santé étant à la fois fortement différenciées selon le critère de la discipline exercée et hiérarchisées à partir de ce même critère, il nous semble opportun de questionner les ambitions des SIP labellisés. Le partage d'un logiciel

permet-il le partage de l'information ? Les données numériques — informatisées et chiffrées — et leurs usages sont-ils communs à tous les soignants-utilisateurs ?

- 6 Pour répondre à ces interrogations, nous nous saisissons de l'action publique comme instrument « par le haut » de la recherche scientifique « par le bas ». Nous nous demanderons dans quelle mesure les SIP en MSP permettent un usage renouvelé des données de santé. L'examen des matériaux recueillis (encadré 1) permettra la construction d'une analyse critique sur les données, leurs usages et leurs valeurs.

Encadré 1. Méthode et terrains

Les résultats présentés sont le fruit d'une recherche doctorale conduite entre 2021 et 2024. Ils ont été produits par l'analyse de matériaux recueillis par entretiens (semi-directifs et contextuels³), observations in situ et recherches documentaires. Le corpus de littérature grise, principalement composé de documents institutionnels, archivés en ligne et en accès libre, a joué un rôle central dans cette enquête. Il nous a donné accès, dix ans après la conception du label, à la liste des membres du groupe de travail mobilisé pour la rédaction du cahier des charges (ASIP Santé, 2016a), au cahier des charges en question (ASIP Santé, 2016b) ainsi qu'aux usages attendus du SIP (Ministère des Affaires sociales, de la Santé et des droits des femmes, 2015 ; Ministère des Solidarités et de la Santé et Ministère de l'Action et des Comptes publics, 2017). Nous avons ensuite pu confronter les attendus sociotechniques des SIP labellisés à leurs usages effectifs en interrogeant et/ou en observant le travail des équipes pluridisciplinaires. Du fait, notamment, des insatisfactions relatives aux logiciels prescrits, certains soignants n'ont été enquêtés que par entretien. D'autres équipes de soins ont quant à elles accepté notre venue sur le site et l'interrogation de plusieurs de leurs membres — c'est le cas des maisons de santé françaises que nous avons renommées Nord, Est, Sud-Est et Sud. Cette enquête mobilise une cinquantaine d'entretiens réalisés auprès de soignants (médecins généralistes, mais aussi pharmaciens ou infirmières par exemple), de coordinatrices et de secrétaires et quatre séquences d'observations sur site (de trois jours en moyenne).

- 7 En portant une attention spécifique à l'instrument qui conditionne la production des données sur les patients (le SIP), aux utilisateurs qui les saisissent (les soignants) et aux institutions qui les interprètent (qu'elles soient politiques, professionnelles ou scientifiques), nous mettrons en évidence que la *datafication* des soins en maisons de santé renvoie à un processus sociotechnique aux nombreuses bifurcations. L'étude des étapes successives de mise en machine, de mise en données et de mise en chiffres des activités réalisées par les équipes pluriprofessionnelles développée ci-après donnera à voir les décalages entre la technologie prescrite et ses usages situés. Si l'imposition d'un logiciel labellisé devait favoriser la coordination, le SIP en question semble finalement l'entraver.

Mettre en machine : les usages fractionnés du système partagé

- 8 L'interrogation des soignants et la déconstruction des principes de conception du label mettent tout d'abord en évidence que les SIP tels qu'ils sont déployés ne satisfont pas

les objectifs de décloisonnement professionnel auxquels ils sont associés. Alors que les logiciels sont censés être *partagés*, leurs usages semblent au contraire *fractionnés*.

L'entrée par les dossiers des patients

- 9 L'analyse des méthodes de construction du label révèle que la coordination numérique des soins a principalement été problématisée par les institutionnels comme par les associations professionnelles autour de l'enjeu du partage des dossiers des patients. En France, comme dans d'autres pays (Hertzum *et al.*, 2022 ; Jensen et Bossen, 2016 ; Sheikh *et al.*, 2014), les dossiers de santé sont présentés comme la pierre angulaire « d'une médecine de parcours décloisonnée et de parcours de soins efficaces » (ASIP Santé, 2016a, p. 5). Ils constituent, selon la Fédération française des maisons et pôles de santé, un enjeu pour la continuité des soins en tant qu'ils « [facilitent] le remplacement, les absences, la gestion des urgences et [doivent] permettre à terme un travail d'objectivation et d'analyse de la patientèle » (Vezinat, 2019, p. 181). En minimisant les pertes d'informations, les redondances et les répétitions dégradées, les technologies numériques participeraient au « bien soigner ».
- 10 Ces capacités de centralisation et de traçabilité sont d'autant plus cruciales en MSP, où les praticiens intervenant auprès d'un même patient exercent des disciplines différentes, les côtoient dans des contextes variés et agissent sur des segments de la prise en charge distincts. Or, l'ambition de ces organisations pluriprofessionnelles est de garantir une distribution du travail non plus divisée, mais coordonnée. Il est par exemple attendu que le médecin diagnostique et prescrive en cabinet, puis que le pharmacien contrôle la délivrance en officine et, enfin, que l'infirmière s'assure de la bonne observance du traitement à domicile. Le travail en équipe de soins se caractérise donc par la réalisation de tâches complexes et entremêlées que les supports informatiques doivent dénouer pour mieux les lier. Regroupés au sein d'un même système d'information, les dossiers informatisés sont censés constituer un espace de réparation des discontinuités écologiques (Joseph, 1994) propres à la prise en charge multidisciplinaire des patients. Ils ont pour objectif de réunir des actions « discrètes » — au sens de la statistique descriptive — au sein d'une même « fonction » de production des soins — pour filer la métaphore mathématique. Cependant, les systèmes labellisés ne semblent pas parvenir à contenir la diversité des données produites par les équipes soignantes.
- 11 Ils apparaissent encadrés dans une conception du soin peu propice au travail pluridisciplinaire. En nous saisissant des « épreuves d'utilisabilité » (*usability trials*) du label comme médium pour révéler les tâches et les usages que le système soutient (Woolgar, 1990, p. 66), nous avons noté que les logiciels certifiés prennent l'exercice de la médecine générale pour référence. La mise en situation imposée par l'Agence du Numérique en Santé pour accorder (ou non) la certification à l'éditeur candidat se fonde sur le schéma du médecin généraliste qui est le médecin traitant du patient, qui paramètre les droits d'accès au dossier informatisé, qui ouvre un formulaire de consultation, qui code la consultation en saisissant le code diagnostic, qui indique les problèmes en cours et qui prévoit une réunion pluriprofessionnelle (ASIP Santé, 2016c). Le scénario conçu insiste donc sur le rôle de coordinateur du médecin traitant (Bras, 2006), qui dirige les parcours de soins primaires, plus que sur la dimension collective des prises en charge, qui est attendue de l'exercice décloisonné en MSP. Ce

développement des systèmes à partir de l'exercice du médecin généraliste explique pour partie que les interfaces produites matérialisent sa « juridiction » (Abbott, *op. cit.*), entendue à la fois comme l'étendue de son pouvoir et son pouvoir de juger.

- 12 Les SIP labellisés sont avant tout un espace d'inscription de la compétence du médecin traitant, car ils s'organisent autour de la prise en charge de la maladie propre à sa profession. Leur structuration matérialise le raisonnement examen clinique-diagnostic-prescription et ne prévoit pas d'organiser les soins qui succèdent à la rédaction de l'ordonnance, c'est-à-dire les actes paramédicaux. Phénomène corollaire, chacune des rubriques dédiées à ces étapes de saisie du travail *médical* renvoie, dans la machine, à des nomenclatures *médicales*. Le champ dédié à l'encodage du diagnostic est la plupart du temps régi par les règles de la Classification internationale des maladies (CIM), un catalogue de pathologies conçu à destination des praticiens responsables de l'identification des maladies et des causes de décès, donc des médecins généralistes et spécialistes.
- 13 Les modalités de partage des dossiers des patients incarnent quant à elles le « pouvoir » du médecin traitant, qui est défini comme l'utilisateur responsable du dévoilement ou du camouflage des informations sur les patients. Le cas d'usage rapporté plus haut mentionne ce rôle spécifique en prévoyant que le SIP permette au médecin généraliste « d'interdire l'accès d'un seul professionnel de santé de la structure au dossier » (ASIP Santé, 2016c, p. 2). Selon la décision du médecin de partager ou de ne pas partager les dossiers, les cloisonnements existants peuvent être dépassés, mais aussi renforcés. En ce sens, le processus de cadrage des besoins sociotechniques relatifs aux dossiers patients partagés marque le glissement d'une volonté de décroisonnement vers le renforcement du pouvoir médical. Les propriétés techniques (rubriques, langages, paramètres d'accès) des logiciels labellisés correspondent aux prérequis numériques de la médecine générale, et non aux besoins de tous les soignants.

La sortie des non-médecins

- 14 L'entrée par les dossiers des médecins traitants a entraîné des problèmes d'utilisabilité et d'ergonomie pour les autres professionnels, paramédicaux notamment. Sous-représentés au sein du groupe de travail qui a conçu le label⁴, ils ne sont pas parvenus à faire traduire leurs pratiques du numérique en fonctionnalités encadrées par la certification « maisons et centres de santé ». Par exemple, alors que le travail des infirmières se déroule la plupart du temps au domicile des patients, les versions mobiles des SIP labellisés ne sont pas développées de façon satisfaisante.

En fait, les éditeurs n'ont pas évalué le fait que, pour les infirmières, il faut que ce soit mobile. Comme elles vont chez le patient 90 % du temps, elles ont besoin d'avoir quelque chose sur place (Adam, ancien membre du groupe de travail ASIP Santé).
- 15 Une telle relégation des besoins des soignants paramédicaux au second plan du label a pour partie été motivée par les représentations de leur exercice que se faisaient les agents et praticiens mobilisés par l'ASIP Santé. D'après notre enquête, comme celle de Michel Naiditch (2020), leurs usages de l'informatique ne sont pas associés à la tenue des dossiers patients, mais réduits à la facturation des soins dispensés.

Les infirmiers, par exemple, utilisent très peu le SIP. Ils utilisent leur logiciel métier qui, en gros, leur sert principalement à faire de la gestion de tournées et de la

facturation. Ils mettent très peu de données dans les dossiers patients, y'a pas de dossiers médicaux (Stéphane, médecin généraliste).

- 16 Allant dans le même sens, les infirmières elles-mêmes distinguent le SIP labellisé de leur logiciel « métier-métier » sur le critère des fonctions de comptabilité et de portabilité que nous avons mentionnées.

En fait, ce qui te manque sur [le SIP], c'est ton logiciel métier-métier : la facturation, l'agenda, une version mobile... (Leïla, infirmière).

- 17 Bien qu'il soit indéniable que les tâches de facturation composent une large partie du travail de machine (Strauss, 1992) infirmier, puisque les premières concernées l'affirment, nous voudrions néanmoins proposer une analyse plus nuancée que celle des médecins postulant que les infirmières « ne tiennent pas leurs dossiers » (Lucas, médecin généraliste). Selon nous, l'exercice paramédical ne se fait pas sans dossiers patients : il se fait par le biais de dossiers patients dont la logique d'organisation déroge au raisonnement examen clinique-diagnostic-prescription, dont nous avons montré la centralité dans la rédaction du label plus haut. Si les SIP charrient une conception « corrective » de la médecine, en s'organisant principalement autour de la logique dominante — mais contestée (Dumesnil, 2011) — du *cure*, l'exercice paramédical relève d'une logique du *care*, qui ne repose pas sur les mêmes pratiques de l'informatique. Comme le fait remarquer Delphine, qui est orthoptiste, les soins qu'elle dispense n'ont pas nécessairement pour objectif la correction du pathologique. Lorsque la guérison est impossible, ils sont parfois même « infinis ».

J'ai aussi des patients sur du très très long cours, avec des pathos un peu bizarres.

[...] J'ai des patients ça fait déjà 3-4 ans que je les vois et dans 3-4 ans ils seront toujours là. Mais, du coup, je n'ai pas de fin (Delphine, orthoptiste).

- 18 Satisfaire la prise en charge des malades plus que de la maladie (Acker, 1997) implique de prendre de la distance par rapport au moment, certes crucial, du diagnostic. Les logiciels paramédicaux rendent compte de cette conception du temps du soin en s'organisant autour de la récurrence des actes et de leur distribution dans le temps (figure 1), plutôt qu'en partant du surgissement de la pathologie.

Figure 1. Logiciel type : exercice infirmier

Source : Maquette réalisée par l'auteure à partir de notes d'entretien contextuel et d'observation.

- 19 À l'inverse, les dossiers médicaux se caractérisent par la succession antéchronologique de comptes-rendus diagnostiques (figure 2). La temporalité de la maladie se substitue à la temporalité du soin, ce qui ne facilite pas l'organisation spatio-temporelle du travail auprès du malade constitutive de l'exercice paramédical.

Figure 2. Logiciel type : exercice de la médecine générale

Source : Maquette réalisée par l'autrice à partir de notes d'entretien contextuel et d'observation.

- 20 En outre, cette incompatibilité épistémo-professionnelle entre les SIP labellisés et les besoins des praticiens non médicaux se double d'une incompatibilité technique. Les systèmes d'information certifiés par l'ASIP Santé sont par exemple incompatibles avec les dispositifs utilisés au quotidien par les orthoptistes (qui analysent les données produites par le biais d'instruments d'*eye tracking*), les masseurs-kinésithérapeutes (qui font de la stimulation par électrodes) ou les podologues (qui recourent à des tapis de marche). Sur le terrain, l'exclusion des besoins des non-médecins au moment de la conception du label s'est traduite par de nombreuses situations de non-utilisation du SIP prescrit.

La fuite des informations

- 21 Alors que les médecins généralistes interrogés utilisent tous le logiciel labellisé comme logiciel métier, c'est-à-dire comme principal outil de suivi, le phénomène inverse s'observe chez les autres praticiens. Qu'ils soient issus d'une discipline médicale ou paramédicale, la totalité des non-médecins généralistes rencontrés (à l'exception d'une sage-femme) ont conservé le logiciel métier avec lequel ils travaillaient avant leur intégration de la MSP. Puisque les fonctionnalités et les dispositifs propres à leurs disciplines ne sont pas encadrés par la certification, la plupart des membres de MSP n'utilisent le SIP labellisé que de façon ponctuelle, et certains n'y sont même jamais connectés. Les informations de suivi que ces soignants produisent ne sont donc pas

transcrites dans le système labellisé. Par conséquent, les données que le logiciel centralise sont avant tout celles produites par les médecins généralistes, qui, eux, l'utilisent comme logiciel métier tout au long de leurs consultations.

Le SIP c'est le truc qu'on va utiliser toute la journée. En tout cas, en tant que médecin, c'est le logiciel sur lequel on a tous nos dossiers médicaux. C'est celui-là par lequel on va prescrire (Émilie, médecin généraliste).

- 22 De cette mainmise « technologique » des médecins sur les dossiers des patients comme objets matériels découle la mise à l'écart des autres professionnels. Ces derniers perçoivent les dossiers de soins comme le pendant numérique de la juridiction des médecins traitants.

Je t'avoue que je continue à dire « leurs » dossiers. Alors c'est peut-être aussi que [le SIP] n'est pas mon logiciel métier. Je pense que si ça l'était, je dirais « mes » dossiers au même titre qu'eux. Mais là, je continue à dire « leurs » dossiers et mes collègues [infirmières] aussi d'ailleurs (Alexandra, infirmière).

- 23 Les médecins généralistes étant les premiers rédacteurs des informations centralisées dans les dossiers partagés, ils sont également considérés comme les « propriétaires » des données qui y figurent. Une telle attribution des dossiers des patients à leurs médecins traitants affecte les pratiques de lecture et d'écriture des membres des MSP. Plutôt que de saisir les informations recueillies au fil de leur exercice, les non-médecins privilégient l'envoi d'un message interne au docteur.

[Les infirmières] écrivent pas mal dans la section [de communication instantanée] et un peu moins sur les observations [dans le dossier patient], mais au moins elles peuvent alerter en me disant « tel patient a pris trois kilos⁵, il ne va pas bien ». Elles nous contactent par là (Clara, médecin généraliste).

- 24 Autrement dit, les infirmières reproduisent la logique de la transmission hospitalière (Grosjean et Lacoste, 1999) et les médecins destinataires sont finalement ceux qui jugent de la pertinence d'une donnée, de son inscription dans le SIP et de son encodage.

Clémence (médecin généraliste) : Moi, en fonction de ce qu'ils me marquent, je fais un copier-coller.

- 25 Du fait de leur position dans le système des professions, ils « trient » les informations transmises. Du fait de leur statut d'utilisateur principal du système d'information, ce sont eux qui réalisent l'inscription de la donnée communiquée dans le dossier du patient. L'arbitrage entre ordre et désordre (Denis et Pontille, 2014) est conduit par les médecins et selon leurs règles d'évaluation de l'information. Le nouvel outil ne crée donc pas une habilitation (Bailly et Boudot-Antoine, 2018) renouvelée des praticiens à écrire dans les dossiers des médecins. Il renforce l'habilitation des médecins à utiliser le système partagé comme un outil de transcription de leur activité.

Mettre en données : un travail de structuration qui ne va pas de soi

- 26 En étudiant à la fois la façon dont ils orientent la configuration du SIP ainsi que les activités par lesquelles se concrétise la mise en données, nous montrerons que certains médecins traitants exploitent cette position d'utilisateur principal du logiciel pour construire des données numériques représentatives de leur activité. Cependant, ce travail de transcription de leurs prises en charge ne va pas de soi.

Le soin en texte libre

- 27 Dans le monde de la santé, les données numériques font l'objet de réceptions parfois hostiles. Ces réceptions considèrent les données numériques comme étant opposées à l'individualité et à l'intimité du patient (Baszanger, 1986) ainsi qu'à l'indépendance des soignants (Hassenteufel, 1997). Selon cette position, elles renvoient à une menace de dévoilement du secret médical et de l'expertise du praticien (Bloy et Rigal, *op. cit.*). Ces craintes concernant la singularité de la relation médecin-malade et l'autonomie de l'exercice clinique libéral nous semblaient d'autant plus mises à mal en MSP que ces établissements regroupent des professionnels de santé libéraux. Contrairement à d'autres types de structures coordonnées, comme les centres de santé, les maisons de santé ne salarient pas les praticiens, qui y exercent en tant qu'indépendants. Les tâches de coordination qu'ils réalisent sont rémunérées en plus des actes de soins qu'ils réalisent usuellement, par le biais de financements additionnels que versent les pouvoirs publics à la structure en fin d'année civile. Par conséquent, nous nous attendions à recueillir des rejets des SIP fondés sur deux types de critiques. Premièrement, celles relatives au passage du colloque singulier au colloque pluriel⁶ (Sarradon-Eck, 2010, p. 266) qu'implique le partage d'informations. Deuxièmement, celles concernant les craintes du recours à la mise en chiffres comme outil de rationalisation des soins — voire « d'industrialisation » de la médecine libérale (Da Silva, 2018). Nous avons cependant dû réviser nos hypothèses.
- 28 En effet, le partage des informations est présenté comme une richesse par la totalité des membres de MSP rencontrés — qu'ils soient médecins ou non-médecins et qu'ils se positionnent comme lecteurs ou scripteurs.
- C'est cadré, c'est joli, c'est propre. Tout le monde peut lire parce que ce n'est pas écrit en pattes de mouche, c'est clair (Carole, médecin généraliste).
- À la [MSP], on a une qualité de la consult' qui est très supérieure. On a quelques médecins [du coin] qui ne font pas partie du pôle de santé, et ceux-là, je ne sais pas comment ils font : il n'y a rien de tracé, rien d'écrit (Catherine, pharmacienne).
- En tant que sage-femme, ces pathologies, j'ai besoin de savoir — que ce soit dans le cadre du suivi gynécologique, pour les IVG ou du suivi de grossesse. Et puis j'ai rajouté mon grain de sel : dans les trucs [les rubriques personnalisées du dossier patient, ndla], il y a aussi mon suivi de grossesse, qui prend un gros pâté. [...] Moi aussi j'ai rajouté des trucs dans leur trame [celle des médecins généralistes, ndla] et personne n'a jamais râlé (Margaux, sage-femme).
- 29 La mise en commun des dossiers est perçue — bien qu'à des degrés divers — comme une occasion d'enrichissement de l'activité, car elle permet de réduire la dépendance au patient et de sécuriser l'acte professionnel (Bloy, 2008).
- 30 De plus, si la menace d'une rationalisation « par le haut » de l'exercice libéral inquiète quelques-uns des praticiens interrogés, elle n'est pas confirmée par la place prise par les données numériques dans les politiques de santé publique considérées. Le recours à des objectifs intermédiaires est en effet fréquent, puisque les soignants n'ont pas les moyens de contrôler le comportement des patients. Le médecin ne peut pas être évalué sur l'amélioration des résultats de biologie d'un patient diabétique, puisqu'il ne peut pas s'assurer qu'il respecte sa diète et son traitement. Les métriques cliniques et thérapeutiques ne pouvant être vérifiées n'ont pas à être enregistrées. Ce raisonnement est particulièrement visible dans le cas des maisons de santé, où ni les usages du SIP ni l'encodage des dossiers patients ne sont imposés puis évalués (Plegat, 2024).

- 31 C'est notamment ce contexte peu coercitif qui explique que les soignants non-médecins n'utilisent pas le logiciel labellisé et que les médecins généralistes, qui sont quant à eux les principaux utilisateurs du système d'information, adoptent principalement des pratiques d'encodage servant à atteindre les objectifs classiques de l'informatique, à savoir l'automatisation et la standardisation (Boullier, 2016). Les principes de structuration des écrits auxquels ils adhèrent concernent l'utilisation des fonctionnalités de génération de « documents intelligents » selon l'expression employée par Nolwenn, représentante commerciale d'un éditeur de logiciel. En saisissant les informations relatives à la santé des patients et à leur situation administrative dans des champs dédiés, ils génèrent automatiquement des « Cerfa⁷ favoris » (comme ceux relatifs aux arrêts de travail ou aux certificats d'aptitude à la pratique du sport), des feuilles de soins (nécessaires au remboursement par la sécurité sociale) ou des « ordonnances type otite ou grippe », pour reprendre les exemples choisis par Alfred, salarié d'un éditeur de logiciel labellisé. Ces fonctionnalités de production des documents médico-administratifs redondants ne suffisent toutefois pas à saisir l'épaisseur du travail de transcription de leurs activités que conduisent certains médecins généralistes.

La médicalisation du fonctionnement mécanique

- 32 Les médecins les plus appétents en termes d'informatique personnalisent les dossiers des patients en y intégrant des protocoles de soins ciblés sur des pathologies spécifiques. Ces formulaires prennent la forme de scripts numériques qui accompagnent l'encodage en contenant les méthodes de rédaction des praticiens. Ils participent à la construction de bases de données standardisées puisqu'ils contribuent à transformer le caractère particulier du colloque singulier en différentes données « neutralisées », au sens d'« objectivées ». Cette objectivation passe à la fois par la structuration du dossier patient en différents champs représentant les critères de sélection des informations pertinentes et par l'emploi de langages médicaux. On retrouve, la plupart du temps, la Classification commune des actes médicaux (CCAM), la Classification internationale des soins primaires (CISP), la nomenclature générale des actes professionnels (NGAP) ou encore la CIM, qui est probablement la plus connue. Grâce à ces langages normés et à ces rubriques prédéfinies, les éléments généraux sur la situation du malade (âge, sexe, antécédents, traitements) prennent la forme de cases à cocher (oui/non, plusieurs modalités), à définir parmi une liste d'options (volet déroulant) ou à remplir (en texte libre).
- 33 Tout standardiser n'est pas pour autant une pratique valorisée. La compétence professionnelle réside dans la distinction entre la normalisation pertinente et celle qui ne l'est pas. Il est par exemple souhaitable d'encoder les informations relatives au suivi de l'hypertension artérielle, car il s'agit d'une maladie chronique qui peut avoir de lourdes répercussions si elle n'est pas traitée (comme un accident vasculaire cérébral ou une crise cardiaque). Il est, à l'inverse, non pertinent de numériser les données liées à une rhinopharyngite, qui est une infection virale bénigne, saisonnière et aux symptômes peu graves (tels que des écoulements du nez). Les choix de *datafication* du soin relèvent donc déjà d'un travail professionnel d'identification des besoins en informations fondé sur la fréquence de confrontation à la pathologie, sa dangerosité ou encore l'évolution des recommandations cliniques, diagnostiques et thérapeutiques.

- 34 En outre, si la « formularisation » de toutes les pathologies n'est pas souhaitable, il en est de même pour la nature des éléments encodés à l'échelle d'une même pathologie, d'une même situation ou d'un même traitement. Les médecins en question cherchent à résoudre les « tensions entre tentative d'exhaustivité et indexicalité des informations [qui s'ancrent] dans des écarts entre des formes de pertinences pratiques incompatibles (Dodier, 1993), par exemple assurer le suivi intelligible d'un patient en particulier versus produire des statistiques agrégées » (Denis et Pontille, *op. cit.*), en intégrant les classifications permises par le SIP aux taxinomies structurant la réflexion en médecine générale. Ils se projettent dans le fonctionnement du logiciel afin de concevoir un système de catégorisation médicale le plus ajusté possible à la catégorisation que permet techniquement le système. Par exemple, pour mettre en œuvre des pratiques d'écriture aux rendements optimisés, les utilisateurs ne créent pas un formulaire pour chacune des pathologies qu'ils ont identifiées comme étant pertinentes. Ils les regroupent par famille selon les symptômes, les traitements ou les comorbidités associés. Les maladies cardio-vasculaires sont ainsi réunies au sein d'un même document, car elles nécessitent toutes de prendre en compte des éléments relatifs au mode de vie des patients (tabagisme, consommation d'alcool, pratique d'une activité physique) et sont traitées par les mêmes familles de médicaments.

L'encodage comme monopole médical

- 35 Pour inscrire une information dans un champ permettant son encodage, l'utilisateur doit ensuite identifier la rubrique et l'expression compatibles avec la nomenclature de saisie prévue. Or, comme nous l'avons vu, les SIP sont avant tout des logiciels médicaux. Les données encodées doivent donc satisfaire le langage du médecin généraliste à la fois dans son contenu (la nature de l'élément) et dans sa forme (le vocabulaire employé). Par conséquent, la « bonne information » comprend un enjeu de « bonne formulation ». Elle doit être pensée selon sa pertinence sur le plan médical et exprimée dans les termes de la médecine générale.

Les bons termes, ce sont leurs termes à eux, médicaux, purs et durs. [...] Le mot qu'on emploie c'est le signe de reconnaissance. Mais c'est le monde médical, c'est comme ça (Leïla, infirmière).

- 36 Dans ce contexte, les praticiens paramédicaux rencontrés, et notamment les infirmières, s'emploient à façonner leur manière de communiquer afin qu'elle se rapproche de celle des médecins généralistes. De la même façon que les médecins légistes traduisent l'autopsie pour faire avancer l'enquête judiciaire (Juston Morival, 2021), les infirmières traduisent leur diagnostic métier pour faire avancer la prise en charge médicale. Elles façonnent leurs écrits pour satisfaire ce référentiel à la fois dans la forme et dans le fond, quitte à vider la retranscription numérique de leur travail de la spécificité de leur métier.

Je vais aller décrire exactement ce que j'ai réalisé comme soins, en mettant des termes plus techniques, mais aussi plus « généralistes » parfois. Parce que les médecins ne sont pas forcément doués en cicatrisation, ils ne comprennent pas forcément le nom des produits qu'on met. Toi, c'est pareil, si je te dis « j'ai mis du Sorbact® », tu ne sais pas ce que c'est du Sorbact®. Les médecins du pôle ne sauront pas ce que c'est du Sorbact®. Par contre, si je mets « hydrocellulaire », c'est un terme générique. C'est comme si je te dis « Kleenex® » et « essuie-main », c'est pareil : si tu leur mets des grands termes génériques, ils vont comprendre, si tu leur mets un nom de marque, ils ne comprendront pas [...] Alors que mes collègues, je

vais être beaucoup plus brève. Le langage, le vocabulaire n'est pas le même (Alexandra, infirmière).

Alors, c'est vrai que ça change notre langage parce que dans mon logiciel de traçabilité métier, j'y vais franco. On passe du temps à faire ces putains de messages quand même, parce qu'on essaie de pas faire de faute, de parler un peu « professionnel » (Aurélié, infirmière).

- 37 Ce retravail du vocabulaire n'est pas neutre. Comme dans le cas des pratiques scripturales des assistantes sociales étudiées par Delphine Serre (2008), les hiérarchies sociales et professionnelles se cumulent. La peur du « mal écrire » concerne à la fois l'erreur *sociolinguistique*, c'est-à-dire l'emploi d'une formulation informelle ou trahissant une mauvaise maîtrise du français écrit, et l'erreur *technique*, c'est-à-dire l'emploi d'un terme médical inapproprié à la situation.

Quand quelqu'un a fait une hospit' pour un problème très précis, un problème médical, on a la chance d'avoir au pôle des médecins qui nous font confiance. Ils savent que si on dit « décompensation cardiaque », je sais ce que c'est. Je vais aller lire le compte-rendu médical, je vais le décrypter, je sais de quoi on parle. Quand je vais rediscuter avec eux par écrit, je vais réutiliser les bons termes normalement. [...] Mais je ne suis pas non plus médecin (Leïla, infirmière).

- 38 Cette transcription met bien en évidence les tensions qu'engendre l'utilisation d'un logiciel médical comme SIP. D'un côté, les informations produites par les infirmières n'y trouvent leur place que si elles sont formulées selon le langage médical. De l'autre, les infirmières ne sont pas totalement autorisées à employer ce référentiel, car elles ne disposent pas du mandat (Hughes, *op. cit.*) professionnel qui y est associé. Elles ne peuvent donc s'approprier (au sens de faire leurs) ni la machine ni ses fonctions de structuration des données. Ces enjeux de transformation langagière ne se limitent pas au métier d'infirmière. Laetitia, psychologue, explique quant à elle ne pas souhaiter alimenter le dossier patient précisément, car elle considère que cela l'obligerait à adopter le vocabulaire du *Diagnostic and Statistical Manual* (DSM). Ce référentiel conçu par l'Association américaine de psychiatrie est le symbole du passage de son cœur de métier, la psychologie, à son équivalent médical, la psychiatrie. Loin d'être anodine, cette traduction donne lieu à des tiraillements.

J'ai pas spécialement envie de marquer [dans les dossiers partagés] parce que, ce qui est utile pour moi, dans mon jargon à moi de psycho, ça veut dire qu'il faut que je le traduise en termes DSM. Donc ça me demande une pratique différente (Laetitia, psychologue).

- 39 Laetitia préfère ne pas l'inscrire dans le SIP plutôt que de le retranscrire dans un autre référentiel, qu'elle perçoit comme une transcription dégradée. Les cas présentés montrent ainsi comment les décalages entre la définition numérique des soins et les conceptions métier de l'activité soignante creusent un écart entre les données encodées et les données exploitées par les professionnels au quotidien.

Mettre en chiffres : une exploitation partielle et partielle des données de santé

- 40 Le SIP ne gardant pas de trace des tâches effectuées par tous les soignants, il semble cohérent d'investiguer quels savoirs ces données partielles font émerger et lesquels elles maintiennent immergés (Vinck, 1997). En analysant les usages qui sont faits des

informations enregistrées dans les SIP, nous proposerons une réflexion sur leur (in)capacité à saisir la prise en charge pluriprofessionnelle.

La valeur scientifique comme privilège médical

41 La mesure étant au cœur du développement de la médecine à la fois comme science et comme profession (Maire et Thalabard, 1994), certains praticiens exploitent les données cliniques qu'ils ont recueillies pour étendre leur expertise au-delà des prises en charge individualisées des patients et en dehors des temps de consultation. Dans les maisons de santé encodant leurs données, les médecins généralistes font du SIP une ressource pour se positionner en acteurs légitimes de la recherche, dont le secteur libéral est usuellement exclu (Bloy et Schweyer, 2010). À partir des informations saisies dans le système labellisé, ils produisent des statistiques inspirées de la démarche quantitativiste conçue par Claude Bernard. Grâce à l'intégration de formulaires numériques et à la standardisation des données qu'ils favorisent, les médecins généralistes développent des usages renouvelés des bases de données. Ils créent par exemple des requêtes dédiées à la réalisation de recherches, d'actions de prévention ou d'analyses de pratiques.

42 La MSP Nord, qui a pour ambition de « devenir une maison de santé pluriprofessionnelle universitaire (MSPU) » (Christophe, médecin généraliste), sur le même format que les centres hospitaliers universitaires, souhaite permettre à l'un de ses étudiants de travailler sur l'apnée du sommeil. Cette équipe réfléchit aussi à une manière d'encoder systématiquement le métier du patient afin de construire une étude sur les maladies professionnelles. La MSP Est, quant à elle, se tient informée des programmes menés par la Haute Autorité de santé (HAS) (une autorité publique à caractère scientifique), confronte les caractéristiques de sa patientèle aux caractéristiques de l'échantillon recherché et conclut de sa participation — ou non — à l'expérimentation en cours.

C'était la HAS et c'était sur l'antibiothérapie dans les infections cutanées. Moi, j'ai été requêter « plaies » et « abcès » dans tous nos motifs de consultation sur deux ans. Et puis j'ai croisé ça avec les codes ATC⁸ des antibiotiques. Je me suis rendu compte qu'il n'y en avait quasiment pas, qu'on n'était pas des prescripteurs d'antibiothérapie locale, ni par voie générale, dans les problèmes d'infections cutanées et donc que je ne pouvais pas prétendre avoir une quelconque expertise sur le sujet. L'idée du groupe de la HAS étant évidemment de produire un référentiel de bonnes pratiques amenant à réduire de manière très importante la prescription d'antibiotiques dans les infections cutanées. Mais on était assez détendu de notre côté puisqu'on s'est rendu compte qu'on n'était pas prescripteurs de ce truc, mais que pour autant, ça ne faisait pas de nous des experts puisqu'on avait simplement appliqué ce qu'on sait depuis très longtemps : quand t'as un furoncle tu mets pas des antibiotiques dessus ! [rires] D'abord, ça ne marche pas, et puis ensuite, tu fais une sélection de germes résistants aux antibiotiques quand tu utilises ces produits. Mais on le sait depuis très longtemps, on a juste appliqué ! (Éric, médecin généraliste)

43 Le médecin généraliste — qui est aussi le cofondateur et le gérant de la MSP — explique que, la HAS cherchant à limiter la prescription d'antibiotiques dans les cas d'infections cutanées, ses collègues et lui ont construit une requête combinant les champs de saisie « comptes-rendus de consultations → plaies et abcès » et « traitements prescrits → antibiotiques » pour objectiver le fait que les médecins de la MSP ne sont pas des

prescripteurs d'antibiothérapie dans le cas d'infections de la peau. En interrogeant le SIP, il atteste du fait que son équipe respecte les bonnes pratiques de prescriptions médicamenteuses souvent négligées ou oubliées (Mousquès *et al.*, 2001) auprès de la HAS. Or, cet organe est perçu comme une organisation légitime, car menant des activités scientifiques, et comme une ressource pour l'extension de la juridiction des médecins, car son fonctionnement est motivé par la triple ouverture des sphères administrative, scientifique et médicale (Benamouzig et Besançon, 2005).

- 44 Ce partenariat est notamment rendu possible par la conception partagée que se font les médecins militant pour l'encodage des données en MSP et la HAS de la « bonne démarche clinique ». Qu'ils soient libéraux ou institutionnels, ces acteurs valorisent l'objectivation chiffrée des pratiques cliniques (Dodier, 2007) et l'évaluation de leur qualité au regard des preuves externes disponibles (Sackett *et al.*, 1996). Par un processus analogue aux essais contrôlés randomisés caractéristiques de l'*Evidence Based Medicine* (EBM), qui émergent dans le but de trouver les meilleurs traitements possibles dans une situation donnée (Marks et Bouillot, 1999), les requêtes s'apparentent à des « essais contrôlés randomisés » appliqués aux pratiques. Ici, les médecins cherchent à appliquer les méthodes de la recherche en sciences médicales non pas pour proposer de nouveaux traitements, comme dans le cas de l'EBM (Steichen, 2018), mais pour objectiver leur respect de recommandations dont l'efficacité a déjà été prouvée. Ils requêtent pour (faire) savoir comment ils traitent. Les généralistes mobilisent leur relation favorable aux pouvoirs publics (Vezinat, 2020) ainsi que la valeur dont bénéficient les données numériques — chiffrées et informatisées — pour faire « remonter » leurs savoirs du bas du terrain vers le haut des institutions.
- 45 Cependant, bien que nous assistions à un phénomène de récupération professionnelle de l'évaluation quantifiée, celui-ci ne se déploie pas uniformément au sein de la MSP. Si cette médecine fondée sur les preuves « bricolée » (Meyer, 2012) rappelle le principe établi par Jerome Ravetz selon lequel « la recherche scientifique est un artisanat », elle invite à en examiner un second : « la spécificité du savoir scientifique établi s'explique par la complexité des processus sociaux de sélection et de transformation des résultats de la recherche » (Ravetz, 1971, p. 71-72). Nous le verrons ci-dessous, la transformation des données du SIP en ressources pour la revendication de l'expertise fait partie de ces « processus sociaux complexes » qui sont inégalement distribués selon la profession exercée.

La valeur administrative comme spécificité paramédicale ?

- 46 Nous l'avons mentionné pour analyser l'exclusion des non-médecins de la conception du label : les pratiques de l'écrit des soignants paramédicaux sont souvent décrites comme faiblement standardisées. En entretien comme dans la littérature, les représentations sur l'exercice infirmier s'accordent pour considérer qu'alors que les médecins « codent », pour produire des données expertes, les infirmières « cotent », pour facturer les soins qu'elles dispensent. Là où les professions médicales sont associées au raisonnement scientifique, à l'objectivité ou encore aux méthodes quantitatives (Becker *et al.*, 1971/1977 ; Fox, 1957), les professions paramédicales sont plutôt associées à la démarche expérientielle, à la subjectivité et à des compétences relationnelles (Chauvenet, 1974 ; Kergoat *et al.*, 1992).

- 47 Si ce postulat est pour partie vrai, puisque plusieurs infirmières affirment que « [dans leurs études] on leur parle pas du tout de ce qu'il faudrait qu'[elles encodent] comme données » (Maëly, infirmière), nous souhaitons néanmoins insister sur deux points. D'abord, partir de l'hypothèse d'un « manque d'intérêt évident » (Januel, 2021, p. 7) des paramédicaux pour l'encodage nous semble réducteur sous divers aspects. D'une part, ce postulat gomme les mécanismes de domination matérielle et symbolique entre les disciplines de la santé que cette étude cherche à révéler. D'autre part, il participe lui-même à légitimer ces rapports hiérarchiques. Il suggère que, là où les médecins produisent des informations savantes destinées à être analysées à partir de connaissances techniques valorisables, les infirmières produisent des inscriptions comptables exigées par un acteur peu reconnu, l'assurance maladie (Saint-Lary, Plu et Naiditch, 2011). La description du suivi paramédical comme d'un suivi comptable imposé renvoie donc au bas de la hiérarchie des valeurs en vigueur dans le monde de la santé.
- 48 Pourtant, les docteurs, eux aussi, facturent. La question de la cotation est au cœur des revendications des médecins libéraux depuis plus d'un siècle (Hassenteufel, *op. cit.* ; Steffen, 1987). Plus que cela, elle constitue même l'élément déclencheur de leurs pratiques d'encodage. Ces derniers ont été invités à saisir leurs données de manière structurée à la suite du plan Juppé de 1996, les réformateurs souhaitant « disposer de chiffres permettant de savoir qui fait quoi, pour quelle raison, dans quel lieu, avec quels effets et à quel coût » (Ogien, 2000, p. 292). La structuration de leurs dossiers n'est donc pas uniquement et intrinsèquement médicale : elle interagit avec un raisonnement gestionnaire lisible au prisme de l'exercice libéral. Toutefois, du fait de la position des médecins dans le système des professions, leur portrait semble échapper à cette réduction. À la différence des soignants paramédicaux, leurs pratiques de cotation ne sont pas uniquement ramenées à des pratiques administratives. En ce sens, la réécriture de l'histoire de l'informatisation des médecins confirme que le profil socioprofessionnel de celui ou celle qui utilise l'instrument affecte la valeur attribuée à la mesure produite — comme le relève aussi Olivier Martin à propos de la passation des tests de Hamilton ou de Binet-Simon (2020, p. 154). Pour une partie du monde médical, l'encodage du docteur serait un acte de science, tandis que celui de l'infirmière serait un acte comptable.
- 49 En outre, les pratiques de (non-)cotation des soignants paramédicaux s'expliquent par le fait que leur travail s'inscrit dans une forte dépendance légale et professionnelle (Feyfant, 2017). Leurs activités sont définies par d'autres acteurs qu'eux-mêmes par le biais des prescriptions de soins et médicamenteuses. L'analyse comparative des métiers d'infirmière, de médecin généraliste et de kinésithérapeute menée par Françoise Bouchayer montre par ailleurs que cette subordination se retrouve dans leurs rapports à la quantification. Alors que les médecins prennent contact avec les caisses d'assurance maladie pour « obtenir des clarifications sur certaines caractéristiques de leur profil d'activité », les infirmières et les kinésithérapeutes évoquent davantage « les courriers d'avertissements reçus au sujet d'un patient ou d'une série de patients » (Bouchayer, 2010, p. 183). Ses résultats soulignent que les usages des données d'activité varient d'une discipline à l'autre et que les réactions des soignants face à leur exploitation ne sont pas les mêmes. Les uns se renseignent sur leurs pratiques, les autres les justifient. Dans notre enquête aussi, les infirmières expliquent que leurs pratiques de suivi informatisé des patients sont modelées par la « peur du gendarme ».

Au départ, c'est la peur de mal faire. Au départ, c'est vraiment ça : de venir pour aider quelqu'un et d'être délégué dans mes soins. Ça, c'est un truc qui me hante depuis l'école. Et après, la peur du gendarme. [...] Régulièrement, ils nous traitent de voleurs. Quand c'est pas eux, c'est la télé. [...] En fait, ils font des contrôles aléatoires. Demain, ils me disent, « vous avez coté des pansements complexes il y a deux ans, vous avez fait des conneries pour Madame Tartampion. Comment vous justifiez ça ? ». Il faut que je ressorte le dossier de la dame. Il faut que je ressorte mon bilan initial de plaie, il faut que je montre les photos que j'ai faites tous les 15 jours, toutes les 3 semaines, tous les jours des fois (Leïla, infirmière).

- 50 Les pratiques métriques et informatiques décrites marquent des tensions entre l'évaluation pour connaître et l'évaluation pour surveiller ainsi qu'entre l'inscription pour protéger le patient et l'inscription pour se protéger. Leur analyse apporte un éclairage sur un phénomène plus complexe qu'un « manque de culture de la donnée » (Pierre-Louis, coordinateur de MSP et ancien membre du groupe de travail ASIP Santé). Les infirmières ne mettent pas leur expertise en machine comme le font les médecins, car elles ne perçoivent pas la *datafication* de leur activité comme une ressource pour la visibilisation de leurs qualités professionnelles, et ce, alors que la mise en chiffres occupe une place cruciale dans le processus de reconnaissance du travail (Evetts, 2003) et a fortiori du travail clinique.

La valeur pratique face à la valeur clinique

- 51 Le cas significatif de la réduction des données infirmières à leur valeur légale et économique peut être élargi au postulat de la domination du paradigme médical. Alors que la diversification des praticiens gravitant autour du patient est présentée par les acteurs publics et les membres des MSP comme une ressource indispensable au bien soigner, les apports de la pluriprofessionnalité semblent gommés par les nomenclatures médicales. En effet, de nombreux soignants — y compris les médecins traitants eux-mêmes — jugent la relation médecin-malade comme productrice d'informations biaisées et amputées du fait de la verticalité qui la caractérise (Castel, 2005). Symétriquement, les autres professionnels se positionnent comme le moyen d'accéder à des « informations cruciales dont [les médecins n'ont] pas idée » (Lise, médecin généraliste).
- 52 En côtoyant les patients dans des situations variées, le collectif multidisciplinaire accède à des données complémentaires à celles recueillies en consultation qui, une fois partagées entre collègues, permettent de considérer des symptômes plus psychologiques que physiologiques (comme dans le cas des infections urinaires, qui se manifestent par une irritabilité accrue), de préciser un diagnostic en incluant des éléments propres à l'hygiène de vie constatée du patient (parfois diamétralement opposée à l'hygiène de vie qu'il a déclarée au praticien) ou encore de distinguer l'échec d'un plan de soins d'un problème d'observance du traitement. Cet argument est par exemple avancé par Paul, pharmacien, qui souligne la façon dont les données partagées constituent une ressource pour l'extension de son territoire pharmaceutique (Nouguez *et al.*, 2022, p. 256).

Quand on délivre l'ordonnance pour le cholestérol, par exemple, [les patients] nous disent « non non, c'est bon ». Une fois : ok. La deuxième fois : ils disent encore « non non, c'est bon ». Donc tu demandes : « Pourquoi vous ne le prenez pas ? [...] Ah ouais ça vous fait des effets secondaires ? Ça vous fait des crampes musculaires ? ».

Et ça il ne le dira pas au médecin. Donc il y a plein de médicaments prescrits qui ne vont pas être pris (Paul, pharmacien).

- 53 Quand un soignant constate un manquement dans la prise d'un traitement ou dans le respect d'une diète spécifique puis transmet ses observations au médecin traitant du patient, il le prévient de résultats de biologie non représentatifs. Ces échanges, qui permettent d'identifier la cause de l'échec du plan de soins, évitent un changement de traitement inutile ou déclenchent une modification du traitement mal supporté par le patient.
- 54 Pourtant, ces transmissions plus ou moins formelles, souvent orales et produites selon des référentiels métiers différents, ne sont pas inscrites dans le SIP par les praticiens non-médecins, qui écrivent rarement dans les dossiers patients. Elles n'y trouveraient par ailleurs pas leur place, puisqu'elles ne peuvent pas être retranscrites en langage médical. L'adoption de la CIM ne permet par exemple pas de classer informatiquement le « rappel du podologue à un patient diabétique » comme un acte médical. Seuls le résultat de consultation (« patient diabétique avec lésions de grades 2 et 3 aux pieds⁹ ») ou la prescription de soins (« suivi des lésions de grades 2 et 3 à réaliser auprès d'un podologue ») peuvent l'être. Il ne reste donc, dans le dossier, qu'une trace sur l'état de santé du malade et non des traces des différentes variables ayant affecté cet état. Le système et les langages utilisés invisibilisent ainsi le travail des non-médecins généralistes qui n'utilisent pas le SIP comme outil métier et qui ne mettent pas leur activité en forme selon les exigences de la médecine générale.
- 55 Du fait de ces nomenclatures organisées autour de la pathologie et de son diagnostic, les données des non-médecins plafonnent au statut de données intermédiaires. Les savoirs légitimes au sein du monde de la santé étant intrinsèquement liés à la pratique de la médecine, les informations produites hors du logiciel médical ou recueillies par des utilisateurs non-médecins sont gommées à mesure qu'elles sont traduites en savoirs médicaux. Dès lors, le travail des données que réalisent les infirmières s'apparente moins à la production de savoirs en tant que tels qu'à des « bribes de sciences » telles que Claire Doquet-Lacoste les a définies, c'est-à-dire comme
- [l]es premières traces, concomitantes à l'observation initiale d'un phénomène qui n'est pas encore pensé comme une découverte et que le chercheur consigne par précaution, avec un langage personnel qui va évoluer, au fur et à mesure de la conceptualisation, se dotant à chaque reformulation d'attributs classificatoires qui correspondent à sa reconnaissance progressive dans le domaine scientifique concerné (2009, p. 8).
- 56 Toutefois, à l'inverse des écrits étudiés par la linguiste, les informations sur la santé des patients (comme le poids ou la tension) ne deviennent des données valorisées que lorsque le médecin généraliste les remodèle à partir de son expertise, les inscrit dans « son » support (le SIP) et les incorpore à son exercice (le suivi d'une pathologie chronique).
- Pour adapter les traitements, on va utiliser les données que les infirmières vont faire remonter. Si elles prennent la tension et tout, le poids, nous, après, on a le suivi et on peut regarder quel poids il faisait tel le jour. C'est intéressant pour le suivi des pathologies chroniques (Carole, médecin généraliste).
- 57 Comme Carole, un grand nombre des médecins interrogés évoquent l'idée de l'infirmière comme relais indispensable à la circulation du « matériau humain » (Goffman, 1968) vers le médecin et donc à la bonne prise en charge. Toutefois, les dispositifs politique et technologique en place ne permettent pas la reconnaissance

formelle de ce rôle professionnel. Malgré l'importance donnée à la prise en compte du contexte social, économique et psychologique dans les politiques de prévention, le rôle que jouent les non-médecins dans le recueil et l'analyse de ces éléments ni totalement médicaux ni totalement étrangers à la santé, ne semble pas explicité et codifié. Il s'en trouve donc invisibilisé. La paternité de la donnée est attribuée au médecin qui l'inscrit comme diagnostic dans le système, et ce, même si elle est le produit de l'interaction entre plusieurs acteurs et différents artefacts.

Conclusion

- 58 Notre recherche s'inscrit donc modestement dans la suite des travaux qui ont montré les effets des hiérarchies professionnelles (Latour et Woolgar, 2006/2013 ; Star et Griesemer, 1989) et de la force des instruments (Clarke et Fujimura, 1996 ; Fujimura, 1996 ; Ravetz, *op. cit.*) dans la production des « bons » savoirs. Malgré la promotion officielle du regroupement pluriprofessionnel, les SIP labellisés prescrits aux équipes multidisciplinaires constituent moins une nouvelle arène à conquérir qu'un espace de maintien de l'autorité professionnelle du médecin traitant dans le champ des soins primaires. Les logiciels ne facilitent pas le partage d'information, ne favorisent pas l'émergence de pratiques de l'informatique similaires et n'autorisent pas des usages communs à tous les soignants, car ils embarquent une conception curative de la santé.
- 59 Au moment de la définition des prérequis du label, les représentations sur les données produites par les soignants paramédicaux se sont heurtées aux conceptions de la donnée légitime. Décrites en termes de facturation, leurs activités de mise en machine et de mise en chiffres n'ont pas été transposées dans les SIP, qui transcrivent les activités médicales au sens strict. Le travail « autour des bases de données » (Dagiral et Peerbaye, 2012) réalisé par ces non-médecins étant encadré par d'autres nomenclatures et servant d'autres logiques que celles du diagnostic et de la prescription, il n'est pas traduit en données numériques. Non seulement il n'apparaît pas dans le système prescrit, qui est sous-utilisé (voire non utilisé), mais, de plus, il est gommé par les statistiques produites sur l'exercice en MSP. Alors que certains médecins cherchent à évaluer leur activité afin de faire reconnaître la qualité des soins qu'ils dispensent ou leur participation aux missions de santé publique, les autres praticiens — paramédicaux notamment — peinent à s'approprier de tels usages du numérique.
- 60 Cela ne signifie pas pour autant qu'ils sont tout à fait étrangers à la force des données chiffrées. En entretien, les soignants paramédicaux sont plusieurs à évoquer les possibilités offertes par les SIP en MSP pour valoriser leur exercice. Maély, une infirmière, évoque le recours à de « fausses nomenclatures » permettant de faire de leurs activités une expérience. Elle espère que cette méthode d'objectivation d'activités décisives, mais jusque-là non scrutées par les autorités de santé (comme la préparation d'un pilulier chez une personne âgée polymédicamentée), permettra à son groupe professionnel de se faire entendre¹⁰. En ce sens, notre étude se positionne donc, plus largement, dans la continuité des recherches s'étant attachées à observer les liens entre quantification et professionnalisme — qui structurent les métiers du journalisme (Grossi, 2020), de la politique (Theviot, 2019), de l'éducation (Grant, 2022 ; Vaquero, 2020) ou encore de l'ordre (Boussard, 2010 ; Matelly et Mouhanna, 2007).

BIBLIOGRAPHIE

- ABBOTT Andrew (1988), *The system of professions: An essay on the division of expert labor*, Chicago, The University of Chicago Press.
- ACKER Françoise (1997), « Sortir de l'invisibilité : le cas du travail infirmier », dans Bernard CONEIN et Laurent THÉVENOT (dir.), *Cognition et information en société*, Paris, Éditions de l'École des Hautes Études en Sciences Sociales, p. 65-93.
- ASIP SANTÉ (2016a), « Label "Logiciels maisons et centres de santé" », *Présentation du label : Contexte, finalités, évolution, orientations*, Ministère de la santé et de l'accès aux soins, <https://esante.gouv.fr/produits-services/label-e-sante>, page consultée le 13 mars 2025.
- ASIP SANTÉ (2016b), « Label "Logiciels maisons et centres de santé" », *Référentiel fonctionnel*, Ministère de la santé et de l'accès aux soins, <https://industriels.esante.gouv.fr/produits-et-services/label-e-sante-pour-les-si-des-maisons-et-centres-de-sante>, page consultée le 13 mars 2025.
- ASIP SANTÉ (2016c), « Vérification de la recevabilité de la candidature – Cas d'usage type », Ministère de la santé, des familles, de l'autonomie et des personnes handicapées, <https://industriels.esante.gouv.fr/produits-et-services/label-e-sante-pour-les-si-des-maisons-et-centres-de-sante>, page consultée le 13 mars 2025.
- BAILLY Adrien et Florent BOUDOT-ANTOINE (2018), « Renforcement et transgression du cadre de l'intermédiation numérique », *Réseaux*, 212(6), p. 121-148.
- BASZANGER Isabelle (1986), « Les maladies chroniques et leur ordre négocié », *Revue française de sociologie*, 27(1), p. 3-27.
- BECKER Howard S., Blanche GEER, Everett C. HUGHES et Anselm L. STRAUSS (1961/1977), *Boys in White: Student Culture in Medical School*, Moncton, Transaction Publishers.
- BENAMOUZIG Daniel et Julien BESANÇON (2005), « Administrer un monde incertain : les nouvelles bureaucraties techniques. Le cas des agences sanitaires en France », *Sociologie du travail*, 47(3), p. 301-322.
- BERTILLOT Hugo (2014), *La rationalisation en douceur : Sociologie des indicateurs qualité à l'hôpital*, thèse sous la direction de Christine MUSSELIN, Paris, Institut d'études politiques.
- BLOY Géraldine (2008), « L'incertitude en médecine générale : sources, formes et accommodements possibles », *Sciences sociales et santé*, 26(1), p. 67-91.
- BLOY Géraldine et Laurent RIGAL (2012), « Avec tact et mesure ? Les médecins généralistes français aux prises avec les évaluations chiffrées de leur pratique », *Sociologie du travail*, 54(4), p. 433-456.
- BLOY Géraldine et François-Xavier SCHWEYER (2010), *Singuliers généralistes : Sociologie de la médecine générale*, Rennes, Presses de l'École des Hautes Études en Santé Publique.
- BOSSEN Claus, Lotte Groth JENSEN et Flemming Witt UDSEN (2014), « Boundary-object trimming: On the invisibility of medical secretaries' care of records in healthcare infrastructures », *Computer Supported Cooperative Work*, 23(1), p. 75-110.

BOSSEN Claus, Kathleen H. PINE, Federico CABITZA, Gunnar ELLINGSEN et Enrico Maria PIRAS (2019), « Data work in healthcare : An Introduction », *Health Informatics Journal*, 25(3), p. 465-474.

BOUCHAYER Françoise (2010), « Similitudes et différences des infirmières, kinésithérapeutes, médecins généralistes pour apprécier les similitudes et différences dans l'organisation de la relation professionnelle aux patients », dans Géraldine BLOY et François-Xavier SCHWEYER (dir.), *Singuliers généralistes : Sociologie de la médecine générale*, Rennes, Presses de l'École des Hautes Études en Santé Publique.

BOULLIER Dominique (2016), *Sociologie du numérique*, Paris, Armand Colin.

BOUSSARD Valérie (2010), « Les policiers de la sécurité publique et l'obligation de faire du chiffre » dans Valérie BOUSSARD, Didier DEMAZIÈRE et Philip MILBURN (dirs.), *L'injonction au professionnalisme : analyses d'une dynamique plurielle*, Rennes, Presses Universitaires de Rennes, p. 107-120.

BRAS Pierre-Louis (2006), « Le médecin traitant, raison et déraison d'une politique publique », *Droit social*, p. 59-72.

CASILLI Antonio A. (2019/2021), *En attendant les robots : Enquête sur le travail du clic*, Paris, Éditions Points.

CASTEL Patrick (2005), « Le médecin, son patient et ses pairs : une nouvelle approche de la relation thérapeutique », *Revue française de sociologie*, 46(3), p. 443-469.

CHAMBAUD Laurent et Olivier SCHWARTZ (2016), *La santé publique en question(s)*, Rennes, Presses de l'École des Hautes Études en Santé Publique.

CHAUVENET Antoinette (1974), « L'infirmière hospitalière : Technicienne ou travailleur social », *Projet*, 90, p. 1138-1148.

CLARKE Adele E. et Joan H. FUJIMURA (1996), *La matérialité des sciences : savoir-faire et instruments dans les sciences de la vie*, traduit de l'anglais par Françoise BOUILLOT, Le Plessis-Robinson, Synthélabo.

DA SILVA Nicolas (2017), « Quantifier la qualité des soins, Une critique de la rationalisation de la médecine libérale française », *Revue Française de Socio-Économie*, 19(2), p. 111-130.

DA SILVA Nicolas (2018), « L'industrialisation de la médecine libérale : Une approche par l'Économie des conventions », *Management et Avenir – Santé*, 3(1), p. 13-30.

DAGIRAL Éric, Séverine DESSAJAN, Thomas LEGON, Olivier MARTIN, Anne-Sylvie PHARABOD et Serge PROULX (2019), « Faire place aux chiffres dans l'attention à soi. Une sociologie des pratiques de quantification et d'enregistrement aux différents âges de la vie », *Réseaux*, 216(4), p. 119-156.

DAGIRAL Éric et Ashveen PEERBAYE (2012), « Les mains dans les bases de données. Connaître et faire reconnaître le travail invisible », *Revue d'anthropologie des connaissances*, 6(1), p. 191-216.

DENIS Jérôme (2018), *Le travail invisible des données : Éléments pour une sociologie des infrastructures scripturales*, Paris, Mines ParisTech.

DENIS Jérôme et David PONTILLE (2014), « Une écriture entre ordre et désordre : Le relevé de maintenance comme description normative », *Sociologie du travail*, 56(1), p. 83-102.

DODIER Nicolas (1993), *L'expertise médicale : essai de sociologie sur l'exercice du jugement*, Paris, Éditions Métailié.

- DODIER Nicolas (2007), « Les mutations politiques du monde médical. L'objectivité des spécialistes et l'autonomie des patients », dans Virginie TOURNAY (dir.), *La gouvernance des innovations médicales*, Paris, Presses Universitaires de France, p. 127-154.
- DOQUET-LACOSTE Claire (2009), « Écrits intermédiaires, écritures intermittentes. Carnets, notes, bribes de science », *Langage et société*, 127(1), p. 7-22.
- DUMESNIL Julien (2011), « Problèmes épistémologiques et ontologiques », *Art médical et normalisation du soin*, Paris, Presses Universitaires de France, p. 79-89.
- EVETTS Julia (2003), « The sociological analysis of professionalism: Occupational change in the modern world », *International Sociology*, 18(2), p. 395-415.
- FASSIN Didier et Dominique MEMMI (2004), *Le gouvernement des corps*, Paris, Éditions de l'École des Hautes Études en Sciences Sociales.
- FEYFANT Véronique (2017), « Ce que la dépendance fait au métier d'infirmière libérale. Ethnographie d'une épreuve de professionnalité », *Regards*, 51(1), p. 257-271.
- FOX Renée G. (1957), « Training for uncertainty », dans Robert K. MERTON, George G. READER et Patricia KENDALL (dirs.), *The Student-Physician: Introductory Studies in the Sociology of Medical Education*, Cambridge, Harvard University Press.
- FUJIMURA Joan H. (1996), *Crafting Science A Sociohistory of the Quest for the Genetics of Cancer*, Cambridge, Harvard University Press.
- GOFFMAN Erving (1968), *Asiles : Études sur la condition sociale des malades mentaux et autres reclus*, traduit de l'anglais par Liliane LAINÉ, Paris, Les Éditions de Minuit.
- GRANT Lyndsay (2022), « Reconfiguring education through data: How data practices reconfigure teacher professionalism and curriculum », dans Andreas HEPP, Juliane JARKE et Leif KRAMP (dir.), *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, New York, Springer International Publishing, p. 217-241.
- GROSJEAN Michelle et Michelle LACOSTE (1999), « La relève : ethnographie d'une pratique de parole », *Communication et intelligence collective. Le Travail à l'hôpital*, Paris, Presses Universitaires de France, p. 77-122.
- GROSSI Valentina (2020), « L'autonomie journalistique face aux chiffres. L'interprétation des métriques dans deux entreprises de presse », *Sociologies pratiques*, 40(1), p. 25-37.
- HASSENTEUFEL Patrick (1997), *Les médecins face à l'Etat : une comparaison européenne*, Paris, Presses de la Fondation nationale des sciences politiques.
- HEATON Lorna et Serge PROULX (2012), « La construction locale d'une base transnationale de données en botanique. Une mise en lumière du travail des "petites mains" », *Revue d'anthropologie des connaissances*, 6(1), p. 141-162.
- HERTZUM Morten, Gunnar ELLINGSEN et Åsa CAJANDER (2022), « Implementing large-scale electronic health records: Experiences from implementations of epic in Denmark and Finland », *International Journal of Medical Informatics*, 167.
- HUGHES Everett C. (1997), *Le regard sociologique. Essais choisis*, textes rassemblés et présentés par Jean-Michel CHAPOULIE, Paris, Éditions de l'École des Hautes Études en Sciences Sociales.
- JANUEL Jean-Marie (2021), « 11^{ème} révision de la Classification internationale des maladies : Une opportunité pour les soins infirmiers ? », *Recherche en soins infirmiers*, 145(2), p. 5-6.

- JEANNINGROS Hugo (2021), « Capter, quantifier, gouverner. L'assurance comportementale au service de la prévention ? », *Revue Française de Socio-Économie*, 26(1), p. 47-66.
- JENSEN Lotte et Claus BOSSEN (2016), « Distributed plot-making creating overview via paper-based and electronic patient records », *Scandinavian Journal of Information Systems*, 28(2), p. 3-27.
- JOLY Pierre-Benoît (2010), « On the economics of techno-scientific promises », dans Madeleine AKRICH, Yannick BARTHE, Fabian MUNIESA et Philippe MUSTAR (dir.), *Débordements : Mélanges offerts à Michel Callon*, Paris, Presses des Mines, p. 203-221.
- JOSEPH Isaac (1994), « Attention distribuée et attention focalisée : Les protocoles de la coopération au PCC de la ligne A du RER », *Sociologie du travail*, 36(4), p. 563-585.
- JUSTON MORIVAL Romain (2021), *Médecins légistes : Une enquête sociologique*, Paris, Presses de Sciences Po.
- JUVEN Pierre-André (2018), « Des trucs qui rapportent », *Anthropologie et Santé*, 16.
- KERGOAT Danièle, Françoise IMBERT, Hélène LE DOARÉ et Danièle SÉNOTIER (1992), *Les infirmières et leur coordination*, Boulogne, Lamarre.
- LATOUR Bruno (2007), « Le "pédofil" de Boa Vista – montage photo-philosophique », *Petites leçons de sociologie des sciences*, Paris, La Découverte, p. 171-225.
- LATOUR Bruno et Steve WOOLGAR (2006/2013), *La vie de laboratoire : La production des faits scientifiques*, Paris, La Découverte.
- LUPTON Deborah (2017), « Self-tracking, health and medicine », *Health Sociology Review*, 26(1), p. 1-10.
- MAIRE Pascal et Jean-Christophe THALABARD (1994), « Biologie, mesures et médecine(s) », dans Jean-Claude BEAUNE (dir.), *La mesure : Instruments et philosophies* [actes du colloque, Centre d'analyse des formes et systèmes de la Faculté de philosophie de l'Université Jean-Moulin-Lyon III, 28-29 octobre 1993], Seyssel, Champ Vallon.
- MARKS Harry et Françoise BOUILLOT (1999), *La médecine des preuves : Histoire et anthropologie des essais cliniques, 1900-1990*, Le Plessis-Robinson, Synthélabo.
- MARTIN Olivier (2020), *L'empire des chiffres : Une sociologie de la quantification*, Paris, Armand Colin.
- MATELLY Jean-Hugues et Christian MOUHANNA (2007), *Police, des chiffres et des doutes*, Paris, Éditions Michalon.
- MAYER-SCHÖNBERGER Viktor et Kenneth CUKIER (2014), *Big data : La révolution des données est en marche*, traduit de l'anglais par Hayet DHIFALLAH, Paris, Robert Laffont.
- McFALL Liz (2019), « Personalizing solidarity? The role of self-tracking in health insurance pricing », *Economy and Society*, 48(1), p. 52-76.
- MEYER Morgan (2012), « Bricoler, domestiquer et contourner la science : L'essor de la biologie de garage », *Réseaux*, 173-174(3-4), p. 303-328.
- MILLERAND Françoise (2012), « La science en réseau, Les gestionnaires d'information "invisibles" dans la production d'une base de données scientifiques », *Revue d'anthropologie des connaissances*, 6(1), p. 163-190.
- MINISTÈRE DES AFFAIRES SOCIALES, DE LA SANTÉ ET DES DROITS DES FEMMES (2015), « Arrêté du 23 février 2015 portant approbation du règlement arbitral applicable aux structures de santé pluri-professionnelles de proximité », *Journal Officiel de la République Française*.

- MINISTÈRE DES SOLIDARITÉS ET DE LA SANTÉ, MINISTÈRE DE L'ACTION ET DES COMPTES PUBLICS (2017), « Arrêté du 24 juillet 2017 portant approbation de l'accord conventionnel interprofessionnel relatif aux structures de santé pluriprofessionnelles », *Journal Officiel de la République Française*.
- MOUSQUÈS Julien, Thomas RENAUD, Catherine SERMET, Roland CASH et Raphaël SPIRA (2001), « La variabilité des pratiques médicales en médecine générale : Le cas des hyperlipidémies », Centre de recherche d'étude et de documentation en économie de la santé, rapport n°506.
- NAIDITCH Michel (2020), « Difficultés d'usage des logiciels pluriprofessionnels au sein des MSP et mode de régulation du marché des logiciels médicaux : quels liens ? », *Revue française des affaires sociales*, 1(1), p. 237-262.
- NOUGUEZ Étienne, Camille LANCEVELÉE, Delphine MOREAU et Tonya TARTOUR (2022), « Quand le pharmacien s'intéresse à la clinique : Savoirs, coordination et responsabilité dans les activités de prescription », dans Sophie DIVAY, Thomas DENISE, Marie DOS SANTOS, Cécile FOURNIER, Lucile GIRARD et Aymeric LUNEAU (dir.), *Pratiques de coopération en santé, regards sociologiques*, Paris, Association Française de Sociologie, p. 251-261.
- OGIEN Albert (2000), « La volonté de quantifier. Conceptions de la mesure de l'activité médicale », *Annales*, 55(2), p. 283-312.
- PERRIN Raphael, Pierre ROBICQUET et Lucile RUAULT (2025), « Introduction. Où sont les violences médicales ? », *Genèses*, 138(1), p. 3-13.
- PLEGAT Manon (2024), « Les usages de l'informatique en maisons de santé pluriprofessionnelles : de la bureaucratisation contrainte à l'appropriation professionnelle », *Sociologies pratiques*, 48(1), p. 23-37.
- RAVETZ Jerome R. (1971), *Scientific knowledge and its social problems*, Oxford, Clarendon Press.
- SACKETT David L., William M. ROSENBERG, J. A. GRAY, R. Brian HAYNES et W. Scott RICHARDSON (1996), « Evidence based medicine: What it is and what it isn't », *BMJ*, 312(7023), p. 71-72.
- SAINT-LARY Olivier, Isabelle PLU et Michel NAIDITCH (2011), « Adhérer ou pas au CAPI : De quel clivage des généralistes le paiement à la performance est-il le révélateur ? », *Revue française des affaires sociales*, 2, p. 180-209.
- SARRADON-ECK Aline (2010), « Qui mieux que nous ? Les ambivalences du "généraliste-pivot" du système de soins », dans Géraldine BLOY et François-Xavier SCHWEYER (dir.), *Singuliers généralistes : Sociologie de la médecine générale*, Rennes, Presses de l'École des Hautes Études en Santé Publique, p. 253-270.
- SERRE Delphine (2008), « Une écriture sous surveillance : les assistantes sociales et la rédaction du signalement d'enfant en danger », *Langage et société*, 126(4), p. 39-56.
- SHEIKH Aziz, Ashish JHA, Kathrin CRESSWELL, Felix GREAVES et David W. BATES (2014), « Adoption of electronic health records in UK hospitals: Lessons from the USA », *The Lancet*, 384(9937), p. 8-9.
- STAR Susan Leigh et James R. GRIESEMER (1989), « Institutional Ecology, "Translations" and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39 », *Social Studies of Science*, 19(3), p. 387-420.
- STEFFEN Monika (1987), « Les médecins et l'État en France », *Politiques et Management Public*, 5(3), p. 19-39.

STEICHEN Olivier (2018), « Naissance de l'EBM », *De la médecine d'observation à la médecine factuelle : Histoire critique de l'evidence-based medicine*, Paris, Hermann, p. 35-77.

STRAUSS Anselm L. (1992), « La trame de la négociation : Sociologie qualitative et interactionnisme », textes réunis et traduits de l'anglais par Isabelle BASZANGER, Paris, L'Harmattan.

THEVIOT Anaïs (2019), « Les mutations du travail politique au prisme des big data », dans François DUBET (dir.), *Les mutations du travail*, Paris, La Découverte, p. 59-75.

VAQUERO Stéphane (2020), « Utiliser ou contester les "usines à gaz" de l'évaluation par compétences dans les dispositifs pédagogiques ? », *Sociologies pratiques*, 40(1), p. 51-60.

VEZINAT Nadège (2019), *Vers une médecine collaborative, Politique des maisons de santé pluriprofessionnelles en France*, Paris, Presses Universitaires de France.

VEZINAT Nadège (2020), « La promotion conjointe des maisons de santé pluriprofessionnelles : une "communauté d'intérêt" entre association professionnelle et autorités sanitaires », *Revue française des affaires sociales*, 1, p. 79-10.

VINCK Dominique, « La connaissance : ses objets et ses institutions », dans Jean-Marc FOUET (dir.), *Intégration du savoir-faire. Capitalisation des connaissances*, Paris, Hermès, p. 55-91.

WOOLGAR Steve (1990), « Configuring the user: The case of usability trials », *The Sociological Review*, 38(1), p. 58-99.

NOTES

1. Nous employons ici le conditionnel à dessein, le caractère idéalisé de ce portrait type du soignant ayant été nuancé, et contredit, par des recherches récentes (dossier « Violences médicales », *Genèses*, 2025).

2. Jusqu'en 2019, l'Agence du Numérique en Santé s'appelait l'Agence des systèmes d'information partagés de santé (ASIP Santé).

3. À la croisée entre l'entretien et l'observation, l'entretien contextuel consiste à recréer des situations d'utilisation de la technologie afin d'analyser la façon dont l'utilisateur se déplace dans l'interface tout en lui donnant l'occasion de commenter ses actions ou de répondre aux interrogations de l'enquêteur. Cette démarche ethnographique permet non seulement de pallier des accès aux terrains entravés (comme l'impossibilité d'accéder aux consultations médicales), mais également d'objectiver les propos des personnes enquêtées (pour lesquelles il est souvent difficile de verbaliser une difficulté d'utilisation ou un dysfonctionnement par exemple).

4. Sur les 29 membres du groupe de travail ASIP Santé, dont la liste figure dans les annexes du document *Présentation du label : contexte, finalités, évolution, orientations* (ASIP Santé, 2016a), 10 sont des soignants. Parmi eux, nous comptons 6 médecins généralistes, 2 infirmières, 1 chirurgien-dentiste et 1 pharmacien.

5. Les prises de poids soudaines et inexpliquées sont des symptômes d'une insuffisance cardiaque, déclenchant des problèmes de rétention d'eau. Il est donc vital de les prendre en compte et d'en diagnostiquer les causes.

6. En médecine, le colloque « singulier » désigne la relation directe et confidentielle entre un médecin et un patient, tandis que le colloque « pluriel » implique plusieurs professionnels de santé qui échangent ensemble au sujet d'un même patient.

7. En France, cet acronyme de « Centre d'enregistrement et de révision des formulaires administratifs » est utilisé, par métonymie, pour désigner les documents standardisés produits par les administrations publiques.

8. L' *Anatomical Therapeutic Chemical Classification* est une nomenclature de classement des médicaments conçue par l'Organisation mondiale de la Santé.

9. Les patients diabétiques présentent un risque élevé de plaies pouvant conduire à des amputations.

10. De telles initiatives attirent d'autant plus notre attention que les évolutions législatives en cours en France tendent vers une plus grande reconnaissance des soins infirmiers et une formalisation accrue de leurs tâches. La loi du 27 juin 2025, qui fonde juridiquement la consultation infirmière et octroie aux soignantes un droit de prescrire défini sous conditions, impliquera la transformation des nomenclatures en place au moment de notre enquête. Les changements en cours ravivent donc notre intérêt pour les intrications entre cotation des soins, encodage des dossiers informatisés et usages des données numériques en santé et appellent à poursuivre les réflexions qui en ont découlé.

RÉSUMÉS

L'article propose une analyse critique de la *datafication* des soins en maisons de santé pluriprofessionnelles (MSP) portant sur les décalages entre les objectifs de l'informatisation et les usages des données informatisées. S'appuyant sur l'analyse de matériaux qualitatifs (entretiens, observations et littérature grise), il interroge la capacité des technologies déployées en MSP à soutenir la coordination multidisciplinaire, d'une part, et à en visibiliser les effets sur la santé des patients, d'autre part. Les résultats présentés montrent que, malgré la volonté politique de labelliser des logiciels pluriprofessionnels, propices à la centralisation des informations sur les patients, les systèmes d'information examinés soutiennent principalement l'exercice de la médecine générale. Ils ne sont utilisés — sauf à de rares exceptions — que par les médecins généralistes. À l'inverse, ils ne tracent pas les activités des non-médecins, dont le travail est invisibilisé au fil du processus de mise en machine, de mise en données et de mise en chiffres de la coordination pluriprofessionnelle. La *datafication* des soins en maisons de santé relève donc en ce sens d'un pouvoir professionnel inégalement distribué : si certains médecins mobilisent les données informatisées pour faire reconnaître la qualité de leurs pratiques, les autres soignants peinent à s'approprier de tels usages du numérique.

maison de santé pluriprofessionnelle, dossier informatisé, données de santé, professions de santé, coordination

This article offers a critical analysis of the datafication of care in multidisciplinary health centres (MHCs), focusing on the discrepancies between the goals of computerization and the uses of computerized data. Drawing on an analysis of qualitative materials (interviews, observations, and grey literature), the authors examine whether the technologies used in MHCs can support multidisciplinary coordination and make their impacts on patient health visible. The results presented show that despite the political desire to label software as multidisciplinary, which is conducive to centralizing patient information, the information systems examined mainly support the practice of general medicine. With very few exceptions, they are used only by general practitioners. Conversely, they do not track the activities of non-physicians, whose work becomes invisible through the process of mechanizing, datafying, and quantifying multidisciplinary coordination. The datafication of care in health centres thus reflects an uneven distribution of professional power. While some physicians leverage computerized data to have

the quality of their practices recognized, other care providers are struggling to appropriate these digital tools.

El artículo propone un análisis crítico de la *dataficación* de los cuidados en casas de salud pluridisciplinarias referentes al desfase entre los objetivos de informatización y el uso de los datos ya informatizados. Tomando como base para el estudio, los materiales cualitativos (entrevistas, observaciones y literatura gris), se cuestiona, por una parte, la capacidad de las tecnologías implementada en las casa de salud pluridisciplinaria para apoyar la coordinación pluridisciplinaria y por otra parte, hacer visible los efectos en la salud de los pacientes. Los resultados presentados demuestran que, a pesar de la voluntad política de certificar los *software* pluridisciplinarios, propicios para la centralización de las informaciones sobre los pacientes, los sistemas de información analizados sirven de base principalmente para el ejercicio de la medicina general. Se utilizan solamente —salvo en raras excepciones— por la medicina general. Por el contrario, no se marcan las actividades de los que no son médicos, cuyo trabajo es invisible durante el procesamiento de los datos, de identificación digital de la coordinación pluridisciplinaria. La *dataficación* de los cuidados en casa de salud depende, por lo tanto en este sentido, de un poder profesional repartido de manera desigual. Si algunos médicos recurren a datos informatizados para hacer reconocer la calidad de sus prácticas, a las otras personas que también contribuyen en los cuidados se les hace difícil tener acceso a tales herramientas digitales.

INDEX

Keywords : care datafication, multidisciplinary health centre, general practitioners, non-physicians, data uses

Palabras claves : dataficación de los cuidados, casa de salud pluridisciplinaria, médicos generales, los que no son médicos, utilización de datos

AUTEUR

MANON PLEGAT

Manon Plegat est chercheuse postdoctorante, Institut de recherches sociologiques, Faculté des sciences de la société, Initiative média philanthropie, Centre en philanthropie, Université de Genève, Courriel : manon.plegat@gmail.com

Datafication d'un corps faillible et pratiques d'autosoins

Explorer les data valences pour penser la sur-veillance

Sylvie Grosjean, Fabienne Martin-Juchat et Ambre Davat

- 1 Cet article s'inscrit dans la filiation de recherches devenues classiques qui visent à proposer et à tester des modèles d'évaluation des technologies assistées par les usages et les usager·ère·s dans le domaine de la santé. Hélène Prévôt-Huille (2018) a mené un travail bibliographique pour synthétiser les modèles d'évaluation de l'expérience utilisateur et de la satisfaction d'usage. Elle observe qu'au cours des vingt dernières années, des approches anthropologiques ont été intégrées pour compléter celles issues de l'ergonomie et de la psychologie cognitive. En particulier, les travaux de Madeleine Akrich (1991) considèrent que tout dispositif technique fait sens à partir du moment où il réussit à s'intégrer, tout en le re-définissant, au sein du réseau d'interactions signifiantes entre l'usager·ère et son environnement. Au sein de ce courant de recherche, ces dix dernières années, nombre de chercheur·euses ont porté une attention spécifique aux conséquences sociales, culturelles et anthropologiques du processus dit de *datafication* (Ruckenstein et Schüll, 2017 ; Lupton, 2016a ; Smith et Vonthethoff, 2016) dans un contexte de développement de dispositifs de télésuivi.
- 2 Afin d'accompagner les personnes atteintes de maladies chroniques, des technologies (ex. : implants, capteurs corporels, applications mobiles) se sont déployées et jouent un rôle dans la collecte, le traitement et la visualisation des données liées au suivi médical (mesures biométriques, informations sur le comportement ou l'environnement des individus). Or, avec l'essor des maladies chroniques, ces dispositifs techniques intègrent des ambitions d'autosoins (traduction de *self-care* ; Saillant et Gagnon, 1996). Au-delà des fonctions de télésuivi pour les équipes médicales, leur but est de permettre aux patient·es de surveiller eux-mêmes leurs symptômes et de gérer leur traitement et/ou les conséquences de leur maladie de façon partiellement automatisée (Nunes, 2017 ; Pierron, 2020 ; Dagiral *et al.*, 2022). Pour ce faire, elles font appel à divers outils d'auto-surveillance (*self-tracking tools*). Il peut s'agir de simples questionnaires auto-administrés, mais aussi de capteurs corporels dits intelligents, plus ou moins invasifs, et générant « des données de santé » (Grosjean, 2024). Ces données sont traitées par des

algorithmes pour alerter le ou la patient·e et/ou les soignant·es ainsi que pour produire des recommandations ciblées et personnalisées. Dans certains cas, les données de plusieurs patient·es peuvent être agrégées et analysées au moyen des méthodes d'apprentissage machine pour augmenter le niveau de personnalisation des suivis ou recommandations. Les patient·es sont donc encouragés à devenir des acteurs, voire des « gestionnaires » de leur santé en collectant, en surveillant et en interprétant leurs propres données (Prainsack, 2017) au service de différents types activités : alerter en cas d'urgence vitale, suivre l'évolution des symptômes d'une maladie, suggérer des conseils de soins et des ajustements aux traitements, partager des activités de soins avec les aidants naturels, améliorer la collaboration et la communication avec l'équipe de soins ou encore proposer des ressources d'éducation thérapeutique personnalisées (Bol *et al.*, 2020 ; Kreps et Neuhauser, 2013).

- 3 Ces technologies sont ainsi présentées dans leurs promesses comme participant d'une forme d'*encapacitation* des patient·es (traduction de l'*empowerment*), par l'autogestion et l'appropriation compréhensive de leur maladie (Mayère, 2018). De plus, l'usage de ces technologies reconfigure la relation thérapeutique et le « travail du ou de la patient·e » (ou de l'aidant·e). En effet, elles leur confient de nouvelles responsabilités anciennement réservées aux professionnel·les de santé (Mayère, *op. cit.* ; Mathieu-Fritz et Gérard, 2023 ; Oudshoorn, 2008). Ce mouvement de *datafication* de la santé met donc au cœur du processus de soins la responsabilité individuelle, ignorant parfois les inégalités sociales, en excluant ceux qui ne disposent pas tous des mêmes ressources ni des mêmes capacités à gérer les émotions en lien avec la maladie et plus globalement à être « autonome » en matière de santé (Dagiral *et al.*, *op. cit.*).
- 4 Au regard du potentiel de ces technologies et des discours de promesse associés, il convient d'évaluer finement leur plus-value du point de vue du ou de la patient·e, car la qualification de l'expérience corporelle et sensible à une série de chiffres ou de graphiques peut reconfigurer l'expérience de la maladie des patient·es. Dans le cas du diabète de type 1 par exemple, le développement des capteurs de glycémie, puis des pompes à insuline automatisées a radicalement transformé la gestion de la maladie, qui, dans les premiers temps de l'insulinothérapie, reposait essentiellement sur le ressenti des patient·es. Le passage d'une génération de dispositif à l'autre nécessite un réapprentissage, pas seulement sur le plan technique, mais aussi en termes de perception de sensations corporelles (Mathieu-Fritz et Guillot, 2017).
- 5 Ajoutons à cela que l'intégration d'algorithmes permettant la personnalisation du programme de traitement et/ou de réhabilitation par des systèmes automatisés peut affecter la subjectivité par perturbation du rapport du malade à ses sensations corporelles. Cette qualification de la maladie en données peut même engendrer un malaise et un mal-être pour le ou la patient·e (Lupton, 2016b, 2017 ; Brayet, 2022). Plus précisément, la gestion de ces technologies engendre un travail réflexif, en confrontant les patient·es aux « traces » de leurs activités, symptômes de l'évolution corporelle et émotionnelle de leur maladie (Cahour et Licoppe, 2010 ; Lupton, 2017 ; Onno, 2022). Si le but est d'améliorer la santé et la qualité de vie des patient·es, ces technologies sont également susceptibles d'accroître leur détresse, en creusant l'écart entre la prise en compte des symptômes objectivés et le vécu subjectif de la maladie (Lomborg, Langstrup et Andersen, 2020). Enfin, elles s'adressent à une certaine catégorie de patient·es, autonomes et disposant de compétences en matière de littératie en santé et de littératie numérique (Délétroz *et al.*, 2022).

- 6 La question est de savoir si ces formes d'autosurveillance de la maladie amène ou non à transformer la manière dont les individus donnent sens à un corps qu'ils perçoivent comme « faillible » au regard d'une norme imposée par les concepteurs du dispositif de télésuivi. Se joue ici le rapport à la norme imposée par le monde médical comme précédemment souligné dès 1966 par Georges Canguilhem (2013) et le vécu expérientiel de la maladie. Au cœur d'un dialogue entre ressentis corporels et données chiffrées, à quelles conditions un dispositif de télésuivi peut-il accompagner les malades dans leur maîtrise de la pathologie sans faire disparaître leur propre capacité à ressentir et à interpréter un corps qui leur échappe ? Comment une donnée fournie par un dispositif de télésuivi peut-elle favoriser un processus qui fasse sens pour le malade ? Le corps de l'individu souffrant devient un objet de gestion algorithmique et numérique. Par exemple, la douleur devient non seulement une expérience sensible située, mais également une variable à mesurer et à surveiller. N'y a-t-il pas un risque pour l'individu de se distancer de ses sensations corporelles directes, en privilégiant une lecture et une interprétation des données qui sont jugées plus objectives ou précises ? Ou au contraire, dans une certaine mesure, la lecture de données fournie par les dispositifs amorce-t-elle de nouvelles formes de dialogue et de narration entre le ou la patient·e, ce corps vulnérable et l'écosystème de soins (Berson, 2015 ; Brayet et Queval, 2023) permettant ainsi de qualifier ces technologies d'autosoins ? En d'autres termes, le télésuivi peut-il nourrir un processus d'autosoins pour la personne malade ?
- 7 L'objectif de cet article est d'aborder ces questions en tentant de mieux comprendre dans quelle mesure des données de santé¹ — générées par des technologies intégrant des algorithmes — permettent à un ou une patient·e de participer (ou non) à un processus encapacitant, condition *sine qua non* d'une activité ou d'un travail invisible d'autosoin. Cet article vise à dévoiler l'une des dimensions cachées de ce travail du ou de la patient·e impliquant la gestion des données des dispositifs de télésuivi. Pour ce faire, nous nous focaliserons sur les travaux qui ont révélé le travail invisible que les patient·es et soignant·es doivent opérer pour créer du sens à partir des données de santé. Nous nous appuierons plus précisément sur le modèle des *data valences* de Brittany Fiore-Gartland et Gina Neff, qui propose six catégories d'évaluation de la valeur ajoutée des données de santé (Fiore-Silfvast et Neff, *op. cit.* ; Fiore-Gartland et Neff, *op. cit.*). Notre hypothèse dans le cadre de ce projet est la suivante : un dispositif de télésuivi peut être qualifié d'autosoin à partir du moment où les données collectées par celui-ci engendrent de la valeur ajoutée pour les patient·es et les interactions entre patient·es et leur écosystème de soins, valeurs ajoutées qualifiables et évaluables grâce au modèle des *data valences*.

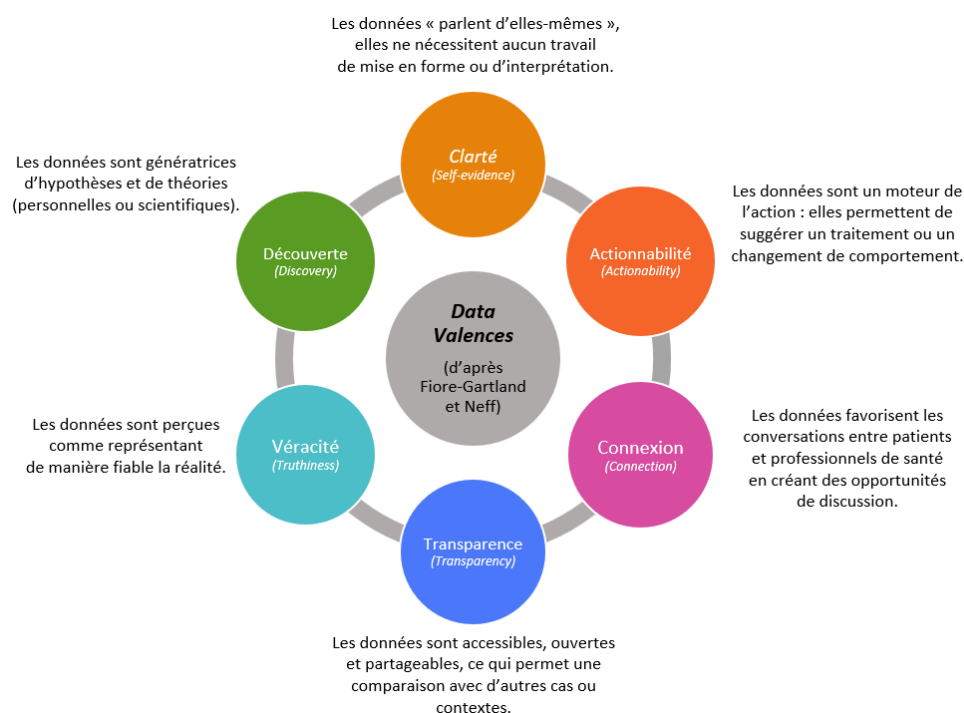
Évaluer le potentiel d'autosoins des dispositifs de télésuivi au moyen du modèle des *data valences*

- 8 Il existe peu de travaux qui permettent d'évaluer la manière dont les dispositifs de télésuivi accompagnent les patient·es dans leurs pratiques d'autosoins. Le modèle des *data valences*, proposé par Fiore-Gartland et Neff (Fiore-Silfvast et Neff, *op. cit.* ; Fiore-Gartland et Neff, *op. cit.*), constitue un cadre analytique permettant d'identifier et de catégoriser les différentes manières dont les données de santé acquièrent de la valeur pour les patient·es et les professionnel·les de santé. Les *data valences* désignent les valeurs, attentes et significations attribuées aux données dans des contextes sociaux,

émotionnels et techniques spécifiques. Cette notion souligne le fait que les données de santé ne sont jamais neutres, elles sont investies de significations multiples selon les acteurs, les usages et les dispositifs dans lesquels elles circulent. Par conséquent, un ou une patient·e va attribuer une valeur à des données de santé — comme preuve, levier d'action, vecteur de connexion, source de découverte, etc. — en fonction de ses besoins, de ses émotions et de sa trajectoire de vie avec la maladie.

- 9 Nous avons choisi d'appliquer le modèle des *data valences* et les six catégories d'évaluation de Fiore-Gartland et Neff sur deux projets de développement de dispositifs de télésuivi (figure 1) : le projet eCARE-PD² au service des patient·es atteints de Parkinson et le projet RealWorld4Clinic³ pour les patient·es atteint·es d'insuffisance cardiaque.

Figure 1. Schéma résumant les six *data valences*



- 10 Source : Schéma créé à partir de Fiore-Gartland et Neff (Fiore-Silfvast et Neff, *op. cit.* ; Fiore-Gartland et Neff, *op. cit.*).
- 11 Dans le cadre de ces deux projets, nous avons mené une étude reposant sur des méthodes qualitatives (ex. : entretiens, groupes de discussion) pour comprendre les attentes des patient·es en matière de valeur ajoutée des dispositifs aux différents stades de développement des projets. Nous avons ensuite catégorisé ces attentes à partir du modèle de Fiore-Gartland et Neff. Ce travail préalable vise à mettre en évidence l'importance du recueil des besoins en matière d'attente de valeur ajoutée. Il souligne également la nécessité d'évaluer le potentiel de valeur ajoutée avec les patient·es le plus tôt en amont, dès les phases initiales de conception (projet RealWorld4Clinic), mais aussi tout au long du processus de conception alternant phases de développement de prototypes et phases d'étude d'usage (projet eCARE-PD). Au plus près des patient·es, ces retours terrains confirment la nécessité de mettre en place un travail spécifique avec des patient·es de coévaluation des outils numériques, dès leur phase de conception,

puis durant les différents stades de développement, quant à leur potentialité encapacitante. La potentialité encapacitante dépend de la capacité des équipes de conception à proposer des données via le dispositif qui puissent faire sens pour les patient·es.

- 12 Il est reconnu dans la littérature que l'usage de technologies numériques de santé implique un « travail d'articulation renouvelé » (Grosjean et Lacoste, 2010, repris par Mayère, *op. cit.*) pour rendre intelligibles et signifiantes des données générées par ces technologies. Afin de mieux comprendre ce travail de construction de sens pour les patient·es, nous mettrons à l'épreuve le modèle des *data valences* (Fiore-Silfvast et Neff, *op. cit.* ; Fiore-Gartland et Neff, *op. cit.* ; Neff *et al.*, 2017). Dans la continuité de grilles d'évaluation de la valeur ajoutée des usages des dispositifs technologiques (dont le plus usité, le modèle UTAU ; Venkatesh *et al.*, 2003), cette conceptualisation permet de mettre en lumière les dimensions émotionnelles, sociales et politiques des données en montrant en quoi elles sont chargées d'attentes (ex. : promesses de personnalisation en santé), d'émotions (ex. : suscitent des réactions affectives) et de tensions (ex. : tensions liées aux multiples interprétations, aux différentes valeurs associées aux données). Dans un contexte de soins, les données ne sont pas uniquement des mesures objectives, mais aussi, par la visualisation, potentiellement vectrices d'émotions positives ou négatives en lien avec la maladie (Verghese, 2008). La *datafication* de sa santé en tant que nouvelle forme d'expérience sensible et affective pose la question suivante : dans quelles conditions des données recueillies et visualisables pour les patient·es peuvent-elles initier des narrations de soi encapacitantes (Nikou *et al.*, 2020) ? Évaluer la dimension encapacitante des technologies s'avère essentiel si l'ambition est de les qualifier de technologies d'autosoins. Pour ce faire, le modèle des *data valences* peut donc éclairer comment les données produites et partagées reconfigurent le rapport des individus à leurs corps, aux symptômes et au vécu de la maladie en nous invitant à explorer les relations sociales et affectives que les individus entretiennent avec ces données de santé. Ce modèle permet de répondre aux questions suivantes : par quels processus des données mises à disposition peuvent-elles engendrer de la valeur ? Peut-on catégoriser plus finement les valeurs ajoutées des données, à savoir leurs capacités à faire sens pour le ou la patient·e ?
- 13 Fiore-Silfvast et Neff (*op. cit.*) parlent de *data valences* (valeurs/valences de la donnée) pour souligner « les attentes multidimensionnelles et les valeurs associées aux données fonctionnant dans différents écosystèmes de données⁴ » (p. 75). Leur modèle, qui repose sur une typologie qu'elles ont développée, permet d'examiner comment les données de santé sont évoquées de manière rhétorique et comment les conversations, les discours, les récits, les pratiques traduisent une variété d'interprétations. Le terme de *donnée* est entendu ici comme un construit sémiotique qui caractérise ce qui est donné à lire, à voir et à entendre pour le malade et les équipes de soins, à partir du dispositif. Dans le cadre de cet article, nous ne détaillons pas tous les enjeux liés à la conception des interfaces de gestion et de visualisation des données précédemment publiées (Davat et Martin-Juchat, *op. cit.* ; Davat, Martin-Juchat et Ménissier, *op. cit.* ; Grosjean *et al.*, 2022). Brittany Fiore-Gartland et Gina Neff nous invitent à porter une attention à la manière dont les différents acteurs narrent leurs relations aux données générées et visualisables par les dispositifs de télésuivi et à explorer les relations à la fois sociales et affectives qu'ils entretiennent avec ces données de santé. Sur la base d'entretiens avec des utilisateur·rice·s, professionnel·les de santé et concepteurs de technologies de santé et bien-être, elles ont identifié six *data valences* différentes (représentées en figure 1).

L'intérêt de ce modèle catégorisant six types de valorisation des données est qu'il permet d'explicitier les attentes et de mettre en évidence les différences de hiérarchisation des priorités non seulement des patient·e·s, mais également de toutes les parties prenantes des projets de télésanté.

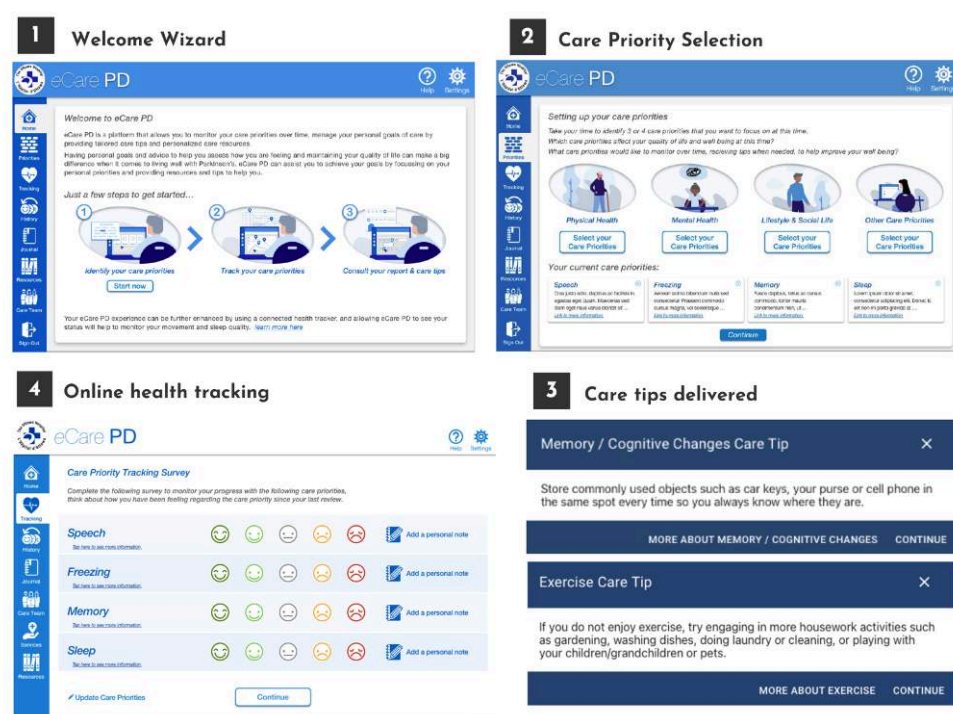
Mobiliser le modèle des *data valences* sur deux projets de co-conception de technologies

- 14 Un dispositif de télésuivi peut être qualifié d'autosoin à partir du moment où les données collectées par celui-ci engendrent de la valeur ajoutée pour les patient·e·s et les interactions entre patient·e·s et leur écosystème de soins, valeurs ajoutées qualifiables et évaluables grâce au modèle des *data valences*. Nous avons mobilisé le modèle des *data valences* sur deux projets de recherche : (a) le projet eCARE-PD, qui implique des personnes vivant avec la maladie de Parkinson et repose sur le développement d'une technologie d'autosoins (*self-care technology*) ayant pour objectif de soutenir les personnes vivant avec la maladie de Parkinson dans leurs pratiques d'autosoins (Grosjean *et al.*, 2022 ; Grosjean *et al.*, 2025) ; (b) le projet RealWorld4Clinic, qui implique des personnes vivant avec une insuffisance cardiaque et accompagne le développement d'un implant de *monitoring* de l'activité cardiaque (Davat et Martin-Juchat, *op. cit.*). Nous avons choisi de mobiliser ces deux projets du fait de leur proximité méthodologique (approche de co-conception incluant un recueil de récits patient·e·s sur leurs usages et leurs attentes) et technologique (technologies qui visent l'autosoin, tout en reposant sur un dispositif de télésuivi alimenté par des algorithmes). Les deux projets sont différents en termes de maturité technologique, ce qui permet de tester le modèle des *data valences* dans deux contextes distincts.
- 15 À partir de transcriptions d'entrevues et de groupes de discussion avec 4 patient·e·s partenaires et 26 patient·e·s, pour le projet RealWorld4Clinic, et de transcriptions d'entrevues avec 18 patient·e·s et 3 groupes de discussion avec des patient·e·s et des aidant·e·s (n=27), pour le projet eCARE-PD, nous analyserons et catégoriserons les attentes et besoins formulés par des patient·e·s en mobilisant la typologie des *data valences* proposée par Fiore-Gartland et Neff (*op. cit.*). Cette typologie permettra d'examiner comment les données de santé sont évoquées de manière rhétorique et comment les conversations, les discours, les pratiques traduisent une variété d'interprétations (Häikiö *et al.*, 2020).

Le projet eCARE-PD et sa méthode

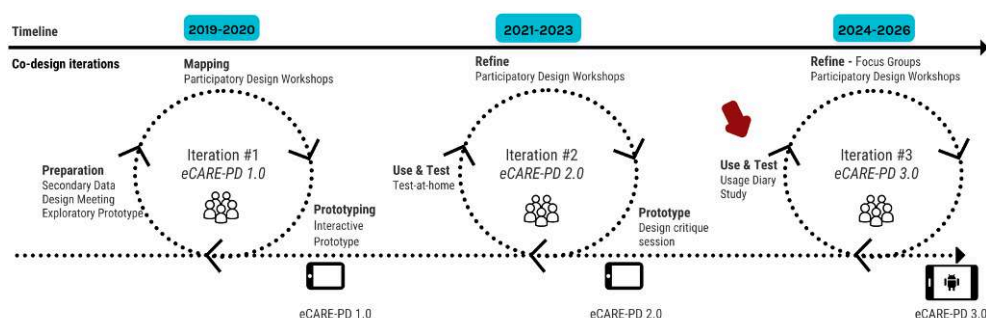
- 16 Le eCARE-PD (*Electronic Care for Parkinson's Disease*) est un outil en ligne qui encourage les patient·e·s à prendre en charge leur maladie en leur fournissant des outils de suivi des symptômes (ou télésuivi), des ressources personnalisées sur la maladie de Parkinson et des conseils de soins (nommés *care tips*). La version 2.0 de eCARE-PD recueille des données sur les priorités des patient·e·s en matière de soins au moyen de divers outils de télésuivi (tels que des auto-évaluations ou des dispositifs portables et capteurs) et analyse les données recueillies pour envoyer des conseils de soins (petits messages sous forme de conseils pratiques) et des liens vers des ressources en fonction du profil de l'utilisateur·rice et de l'évolution de sa maladie (figure 2).

Figure 2. Principales fonctionnalités de eCARE-PD 2.0



- 17 Le projet est né en 2019, avec un premier prototype (eCARE-PD, version 1.0) qui offrait aux personnes atteintes de la maladie de Parkinson la possibilité d'identifier les priorités en matière de soins (ex. : gérer l'anxiété, les troubles du sommeil, la douleur ou faire de l'exercice régulièrement) et d'accéder à une bibliothèque enrichie de conseils éducatifs d'autosoins couvrant les problèmes courants de la maladie de Parkinson et des questions générales liées au bien-être ainsi qu'à des liens vers d'autres ressources recommandées sur le Web. Sur la base d'un premier cycle de co-conception et d'évaluation en situation réelle (figure 3) par un échantillon d'utilisateur·rice·s finaux, certains problèmes clés ont été identifiés et des exigences ont été formulées en vue d'améliorer la plateforme. Les résultats de cette première étude ont été publiés dans Grosjean *et al.* (2022).

Figure 3. Cycles de co-design depuis 2019

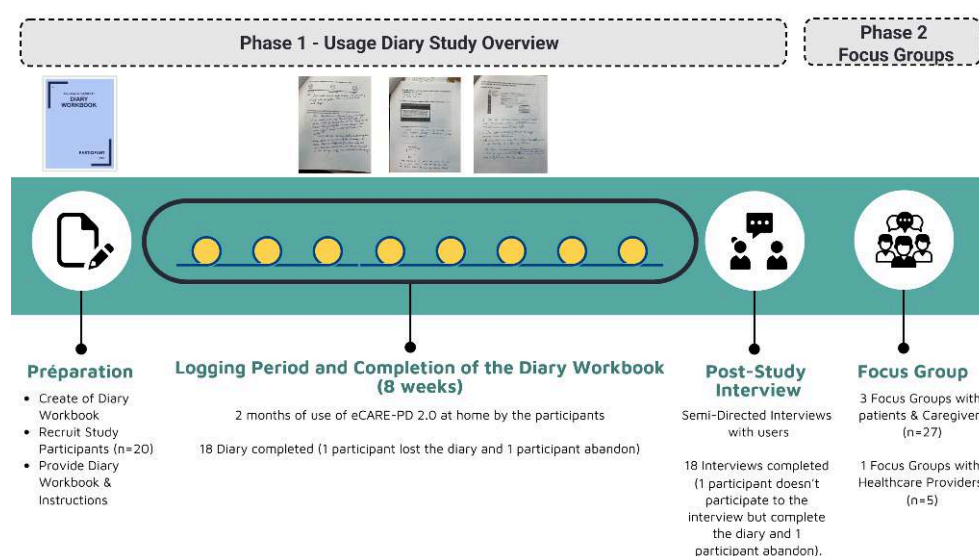


- 18 Une nouvelle version (eCARE-PD 2.0) disponible uniquement en anglais a donc été développée pour intégrer les premiers retours des patient·e·s. Cette version intègre, par exemple, une fonction de géolocalisation (*Care Finder*) afin de faciliter la navigation dans les ressources et services de soins de proximité. Elle fournit des informations sur

la disponibilité et l'accessibilité à ces ressources tant pour les patient·es que leurs aidant·es. eCARE-PD 2.0 offre également la possibilité d'intégrer des données provenant d'appareils portables (ex. : Fitbit) pour un suivi complémentaire de ses priorités de soins ou de ses symptômes. Autrement dit, l'autosurveillance est à la fois autodocumentée (au moyen d'un questionnaire auto-administrée et une fonctionnalité journal) et automatisée (au moyen de la montre connectée).

- 19 Afin d'évaluer eCARE-PD 2.0 avec un échantillon d'utilisateur·rices (n=20), nous avons mobilisé la méthode du journal d'usage (Jarrahi *et al.*, 2021 ; Jokela, Ojala et Olsson, 2015 ; Bardot *et al.*, 2022). L'objectif était de recueillir des données sur le vécu et l'expérience des patient·es lors d'un usage à la maison. Dans cette étude, il a été demandé aux personnes de noter chaque semaine — sur une période de deux mois — leurs activités, interactions, expériences, pensées et détails contextuels lorsqu'elles utilisaient (ou non) eCARE-PD 2.0. Le déroulement de l'étude et la collecte des données ont eu lieu de décembre 2023 à juin 2024 (figure 4). Nous avons jumelé la méthode du journal d'usage (n = 20) avec des entrevues semi-dirigées avec des patient·es et 3 groupes de discussion réunissant des patient·es et des aidant·es. Au début de l'étude, chaque participant·e recevait des instructions, un lien de connexion à eCARE-PD 2.0 et un journal d'usage à remplir chaque semaine pendant une période de huit semaines. Différents types de données ont été recueillis : données démographiques (ex. : âge, stade de la maladie, sexe), descriptions textuelles des activités et des expériences, échantillons de conseils/ressources de soins fournis par eCARE-PD 2.0, réactions et réflexions émotionnelles et retours d'expériences lors des entretiens et groupes de discussion. Au total, 20 participants ont été impliqués dans l'étude qui a généré des journaux d'utilisation hebdomadaires. Des entretiens semi-structurés ont été menés avec 18 patient·es à la fin de l'essai à domicile (n = 18). Les entretiens ont duré de 45 à 75 minutes. Nous avons organisé 3 groupes de discussion réunissant des patient·es et aidant·es (n = 27). Le journal partagé par les participant·es et les enregistrements audio des entretiens et groupes de discussion ont été transcrits pour analyse.

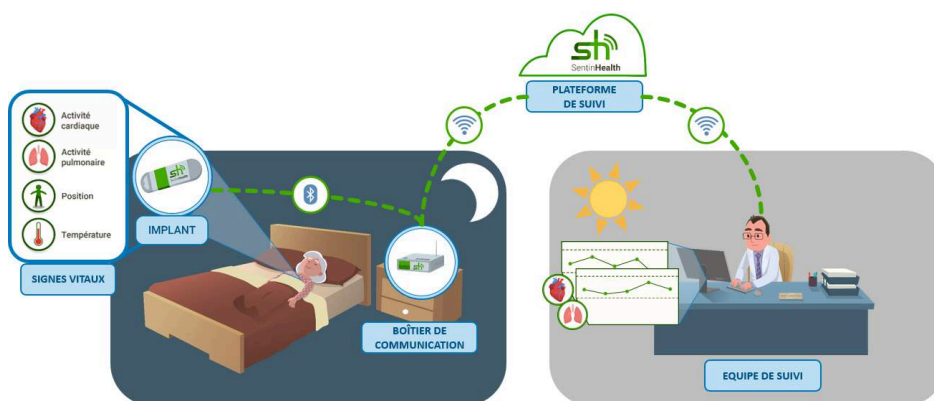
Figure 4. Schéma du déroulement de l'étude de 2023-2024



Le projet RealWorld4Clinic et sa méthode

- 20 Dans le cadre du projet RealWorld4Clinic financé par l'EIT Health⁵, nous avons mené une étude sur les besoins informationnels des patient·es atteints de problèmes cardiaques afin de concevoir une application numérique de télésuivi⁶ d'un implant pour l'insuffisance cardiaque. Le but pour l'équipe de co-conception en sciences humaines et sociales était d'accompagner le développement de l'application numérique de suivi du futur implant de *monitoring* de l'insuffisance cardiaque en s'appuyant sur les attentes informationnelles des patient·es potentiel·les. Il s'agissait notamment de co-concevoir le design de la future application numérique, y compris le choix et le mode des données cardiaques à visualiser (figure 5), destinée aux future·s patient·es. Afin de comprendre leurs besoins informationnels et attentes, nous nous sommes appuyées sur l'expérience de patient·es déjà monitoré·es par différents types de technologies d'autosurveillance et d'autosoins, y compris des montres connectées, des *pacemakers* et des défibrillateurs automatiques implantables.

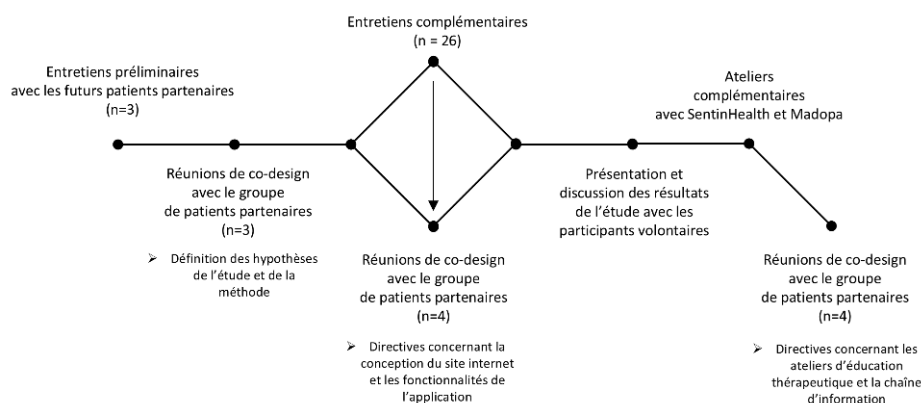
Figure 5. Infographie représentant le dispositif d'implant connecté



- 21 Dans le cadre de ce projet (figure 6), nous avons utilisé une méthode d'entretien qui insiste sur les grandes émotions négatives ou positives du ou de la patient·e en lien avec son environnement, afin de révéler ses attentes. Cette méthode par les récits d'affects a été utilisée dans différents contextes organisationnels (Dumas et Martin-Juchat, 2022, p. 173) et elle s'inspire des travaux de Michael Polanyi (1966) et de Clifford Geertz (1983), entre autres. Elle repose sur une verbalisation de l'expérience par des récits en insistant sur vécu affectif négatif ou positif (l'affect est une catégorie générique incluant les émotions, les sentiments, les passions, les humeurs, etc.), afin de faciliter l'expression des vécus situés et surtout comprendre les systèmes de « valuation » des situations, condition préalable aux processus cognitifs d'évaluation, comme précédemment souligné par George Dewey (2011). Interroger des patient·es en leur demandant d'insister sur leurs émotions permet d'accéder à leur système de hiérarchisation de leurs priorités. Dans le cadre de notre étude, les patient·es étaient invité·es à raconter leur parcours de soins et leur vie quotidienne avec la maladie chronique. Une attention particulière était accordée aux émotions et sensations physiques évoquées ainsi qu'aux relations avec les différents acteurs, à la fois humains (professionnel·les de santé, proches, aidant·es) et non humains (leur environnement composé de médicaments, pèse-personne, applications numériques et dispositifs

médicaux que sont les implants, *pacemaker*, etc.). Pour mener ce travail, un groupe de quatre patient·e·s partenaires a été recruté en mars 2021 pour accompagner le déroulement du projet, à travers l'organisation régulière de réunions de deux heures en visioconférence. Il s'agissait initialement de trois hommes, membres actifs du Résic38, réseau de santé consacré à l'insuffisance cardiaque hébergé dans un CHU (centre hospitalier universitaire). À la suite des premiers entretiens, une quatrième personne s'est jointe au petit groupe : une femme, greffée du cœur.

Figure 6. Schéma récapitulatif du projet et des étapes de recueil des besoins des patient·e·s



- 22 À la suite de cette phase d'identification des besoins énoncés avec les patient·e·s partenaires, 26 entretiens complémentaires ont pu être réalisés entre mars 2021 et juillet 2023 pour affiner les attentes des future·s patient·e·s. Les participant·e·s aux entretiens complémentaires ont été recruté·e·s auprès du Résic38, l'association de patient·e·s souffrant d'insuffisance cardiaque (8 entretiens), d'un hôpital privé français (11 entretiens) et par Internet, du fait de leur implication sur les réseaux sociaux et dans des associations de patient·e·s (7 entretiens). Il s'agit de 16 hommes et 10 femmes, âgés de 21 à 89 ans (65 ans en moyenne). La plupart des participant·e·s (21) étaient déjà équipé·e·s de technologies de *monitoring* implantables (ou l'avaient été par le passé)⁷ : *pacemaker* (5 personnes), défibrillateur (15 personnes) ou *implantable loop recorder* (1 personne). Certaines personnes utilisaient quotidiennement un dispositif d'autosurveillance, constitué d'une application mobile avec questionnaire de symptômes, d'une balance et d'un tensiomètre connectés (12 personnes). Il s'agissait principalement des patient·e·s de l'hôpital privé, qui faisaient partie d'un programme de télésuivi, ainsi que d'une patiente membre d'une association.
- 23 Au total, entre mars 2021 et juillet 2023, 21 réunions de co-design ont été organisées avec les patient·e·s partenaires et deux ateliers ont eu lieu en partenariat avec la *start-up* SentinHealth chargée de commercialiser l'implant et le *Living Lab* Madopa. Ces séances ont permis de concevoir le site Web destiné aux future·s patient·e·s, de proposer un programme d'éducation thérapeutique associé à l'implant et de concevoir un prototype d'application numérique devant permettre aux patient·e·s d'avoir accès à des informations sur leur suivi⁸.
- 24 Chaque réunion de co-design était enregistrée et faisait l'objet d'un compte rendu que les participant·e·s pouvaient amender. Les entretiens étaient enregistrés, puis résumés et transcrits manuellement. Dans la continuité de ce travail de recueil de données, nous

avons extrait des transcriptions issues des entretiens et des séances de travail menés durant ce projet afin de tester la pertinence de la grille d'évaluation par les *data valences*.

Les modalités de la « sur-veillance » comme support à l'encapacitation

- 25 L'usage des technologies de télésuivi s'inscrit dans une dynamique de « sur-veillance » (Dagiral, 2019 ; Lupton, 2016a, 2016b, 2017 ; Onno, *op. cit.* ; Andrieu, 2016), au sens de veille active sur soi, rendue possible par l'intermédiaire d'outils numériques. Nous utiliserons le terme *sur-veillance* afin de souligner la double dimension de toutes pratiques d'autosoins, qui visent à la fois à surveiller (veiller avec autorité, contrôler, observer) et à veiller (faire attention, veiller sur soi). Cette notion ne se limite pas à un simple contrôle ou suivi de ses symptômes, mais appelle une sur-veillance distribuée, car impliquant des acteurs, humains et non humains. Il s'agit alors pour les patient·es de surveiller avec attention leur corps, de veiller sur soi en s'équipant de la technologie (Brayet, *op. cit.* ; Dagiral *et al.*, *op. cit.* ; Lupton, 2016b ; Smith et Vonthethoff, *op. cit.*). La technologie peut devenir un coacteur du suivi médical, produisant des traces qui ont pour but d'aider à interpréter un corps faillible et des symptômes parfois imprévisibles sous condition qu'elle soit conçue à partir du vécu des patient·es. Ces technologies sont pertinentes si elles peuvent assister, voire amplifier la capacité du ou de la patient·e à veiller sur soi, redistribuant ainsi la charge de l'attention entre l'individu, les soignants et la technologie. Dans ce contexte, les technologies — et les outils d'autosuivi qu'elles proposent — engagent un travail d'observation et d'interprétation du corps à travers les données. Les données produites vont acquérir une valeur de preuve qui confère une forme d'autorité aux indicateurs mesurés. L'individu peut alors se reposer sur ces traces numériques pour légitimer un ressenti auprès des soignant·es, comprendre des symptômes imprévisibles et changeants ou ajuster ses routines d'autosoins. L'extrait⁹ suivant est tiré du projet eCARE-PD et il illustre cette activité de sur-veillance médiée par la technologie :

Well, it was interesting to be able to see the progression of like, last week sleeping was here, and this week is better. And so, that brought that into better perspective of what I feel because I realize it's good. Especially something like sleeping, what I thought I was sleeping was not what the system thought I was. It thought I was sleeping more than I felt. So, in that sense it was interesting to see a comparison between sort of actuality and what your mind told you was happening. [...] I would probably say that people would be well to at least look at it and see how they feel themselves because I think that you get lost in what's happening to you today and this is sort of what's happening to you over a whole bunch of days. And I find my symptoms can change from one day to the next. Where, you know, today I'm worried that I'm having trouble walking but tomorrow I'm worried because I'm having trouble making the bed or whatever, you know. And the next day I'm having trouble in the shower. So, I think that that will give people a better overall look at how they're doing overall (Entrevue, Patient 9).

- 26 Dans cet extrait, le patient souligne comment la technologie de suivi lui permet de comparer son ressenti subjectif avec une traçabilité « objective » de ses symptômes. Il réalise, par exemple, que son sommeil perçu ne correspond pas à la mesure du système, ce qui lui permet de se repositionner par rapport à son propre vécu. Cette mise en visibilité des données crée une mémoire du corps, permettant d'identifier des

évolutions qui ne sont pas immédiatement perceptibles au quotidien. Le patient met en évidence une tension entre son ressenti immédiat et ce que la donnée indique. Il est frappé par le fait que le système considère qu'il a mieux dormi qu'il ne le pensait.

- 27 L'un des aspects clés de ce retour patient est la manière dont l'accès aux données permet une prise de recul sur son état de santé. Le patient indique ici que ses symptômes fluctuent d'un jour à l'autre et que, sans cet historique de données, il risque de se perdre dans l'immédiateté de ses sensations du moment. La donnée devient alors un outil de structuration de l'expérience de la maladie. Le patient suggère que la technologie pourrait en aider d'autres à mieux comprendre leur état global et à éviter de se focaliser uniquement sur la souffrance du présent. Cela rejoint l'idée que la surveillance peut être un levier d'encapacitation en permettant aux patient·es de mieux se situer dans leur parcours de soins. Cet extrait montre que la sur-veillance ne fonctionne pas uniquement comme un contrôle externe imposé par la technologie. Elle devient un dialogue entre le vécu subjectif du ou de la patient·e et la donnée, qui peut soit confirmer, soit nuancer son ressenti, et ainsi l'aider à ajuster ses actions et la perception de son corps. C'est ce que nous allons explorer plus en détail en mobilisant notamment la typologie des *data valences* dans les deux projets décrits précédemment.

Une sur-veillance distribuée dans le cadre du projet eCARE-PD

- 28 Dans le cas de eCARE-PD 2.0, la sur-veillance n'est pas uniquement le fait du ou de la patient·e : elle est coproduite par la technologie, qui capte, enregistre et analyse des données, puis les restitue sous une forme interprétable. Le ou la patient·e, pour sa part, interagit avec ces données, compare son ressenti avec les indicateurs fournis et ajuste sa perception de son état de santé. Dans ce cas, la technologie joue des rôles complémentaires dans cette sur-veillance distribuée : (a) elle structure une forme de mémoire de la maladie, permettant de sortir de l'instantanéité des sensations et d'adopter une vue plus large sur l'évolution des symptômes, et (b) elle introduit une comparaison entre perception et mesure, suscitant un travail réflexif chez le ou la patient·e sur son propre ressenti. L'autosuivi de la maladie n'est alors plus une simple observation introspective, mais un processus distribué entre l'individu, la technologie et potentiellement d'autres acteurs (médecins, proches, soignants). Le modèle de *data valences* permet alors d'analyser finement cette dynamique en révélant les différentes manières dont les données de santé « font sens » (ou non) pour les patient·es. Chaque catégorie — preuve, actionnabilité, connexion, etc. — correspond alors à une fonction que les données peuvent remplir dans cette sur-veillance distribuée.

Les données comme preuve de son état de santé (self-evidence)

- 29 Parler de *self-evidence* renvoie au rôle des données comme preuve visible et objective de l'état d'un corps vulnérable, permettant d'attester d'une évolution de la maladie ou de légitimer un ressenti subjectif. Dans le contexte de la maladie de Parkinson, où les symptômes varient quotidiennement et sont parfois difficiles à mesurer précisément, les patient·es utilisent les données générées par eCARE-PD 2.0 pour valider leurs perceptions. Les données sont perçues comme des preuves validant des ressentis corporels ou confirmant leurs sensations subjectives (fatigue, douleur, évolution des symptômes). Cette validation rétrospective permet d'objectiver l'expérience corporelle et de reconnaître les évolutions dans leur état de santé :

Si je trouve que j'ai beaucoup de douleur ou que j'ai beaucoup de stiffness, je l'ai là, peut-être là, je l'utilise une semaine, pour être sûr, pour voir si c'est vraiment pire, ou bien si c'est juste moi qui m'imagine (Entrevue, Patient 5).

- 30 Les participant·es évoquent comment le fait de consulter l'historique de ses données d'autosuiivi constitue une sorte de mémoire objective qui valide ses ressentis et offre une base pour prendre des décisions :

I tracked my history and determined it was more variable than I thought. I struggle trying to decide which each symptom requires attention (Extrait du journal d'usage, Patient 3).

The tracking is what I was using more than anything..., especially when I can combine history with tracking, I could see the variability of my symptoms (Entrevue, Patient 4).

By using the tracking, I realized I had pain (tightness) every day and did not realize before how much tightness (Entrevue, Patient 7).

- 31 L'objectivation des sensations corporelles par les données joue un rôle essentiel dans la perception de la maladie de Parkinson par les patient·es. Le besoin de preuve est particulièrement important lorsque les symptômes sont fluctuants ou ambigus : les symptômes sont-ils liés à la maladie ou seulement liés à l'âge ? Nombre de patient·es se posent la question. Les graphiques et historiques fournis par eCARE-PD 2.0 leur permettent d'associer des chiffres à leurs sensations. Cette confrontation aux traces, ces médiations visuelles les projettent dans un regard rétrospectif et réflexif sur eux (Ruckenstein, 2014).
- 32 Les données offrent donc une base pour se prouver à soi-même et aux autres que les sensations, les changements, les ressentis sont réels et évolutifs. Toutefois, cette objectivation peut engendrer une forme de dépendance aux chiffres qui risquerait d'amplifier le doute si les données ne concordent pas entièrement avec l'expérience subjective. C'est donc un usage ambivalent qui considère les données, d'un côté, comme renforçant la confiance dans le suivi de la maladie et, de l'autre, comme pouvant créer une tension si l'individu perçoit que le corps vécu n'est pas complètement retranscrit par les mesures numériques.
- 33 D'autres chercheur·euse·s ont abordé cette problématique. Par exemple, Lupton (2016a) explique que la quantification du corps peut procurer un sentiment d'encapacitation en offrant des repères objectifs et mesurables. Les données fournies par des technologies de santé permettent aux individus d'objectiver leurs sensations — d'où la dimension *self-evidence* — et de valider des ressentis souvent difficiles à communiquer. Pourtant, cette objectivation n'est pas sans risque. Lorsque les chiffres et graphiques ne reflètent pas entièrement la richesse des expériences vécues, le ou la patient·e peut ressentir une tension entre son corps vécu et son corps mesuré.
- 34 On pourrait s'interroger sur le fait que l'usage de technologies d'autosoins peut amener les patient·es à une moindre sensibilité aux signaux que leur envoie leur corps. Autrement dit, en se reposant de plus en plus sur des données numériques pour comprendre et gérer leur santé, les individus risqueraient de se détacher de leur propre expérience vécue. Concrètement, au lieu de ressentir directement les sensations — douleur, fatigue, variations subtiles de leur état —, les personnes commencent à s'appuyer exclusivement sur des indicateurs chiffrés (comme les mesures de sommeil ou d'activité physique) pour interpréter leur corps. Ce phénomène peut engendrer une transformation du rapport au corps, jadis vécu de manière immédiate et subjective, et progressivement transformé en un objet de suivi.

Les données comme levier d'action (actionability)

- 35 L'actionnabilité des données renvoie à leur capacité à inspirer ou à justifier des changements dans les pratiques d'autosoins. Les informations affichées sous forme de chiffres (historique de l'autosui de ses symptômes) ou de texte (conseils de santé adaptés au profil de l'utilisateur) par *eCARE-PD 2.0* incitent les patient·es à adopter des routines, à modifier leurs habitudes et à s'engager dans une autoprise en charge de leur santé. Par exemple, la plateforme leur envoie des messages personnalisés que les patient·es ont commentés dans leur journal d'usage :

It is obvious from the tip I received that I didn't have enough medication before going to sleep at night (Extrait du journal d'usage, Patient 1).

Decided to look at some resources — Interaction food + Ledopa — carbonated drink with pills (I do this because I dislike water therefore use soda stream to carbonate water daily), happy to learn beneficial to do this (Extrait du journal d'usage, Patient 2).

- 36 Ces extraits montrent que les patient·es exploitent les informations générées par *eCARE-PD 2.0* pour ajuster leur routine de soins ou pour chercher en apprendre plus sur leur traitement, leurs symptômes. Plusieurs participant·es ont souligné que l'accès à des visualisations et tendances dans *eCARE-PD 2.0* les incitait à revoir certains comportements ou à en adapter d'autres. Par exemple, un participant souligne que l'accès à une vue hebdomadaire de son suivi lui permet de détecter des variations et de mieux comprendre certains changements ou fluctuations dans ses symptômes :

started doing it just once a week. [...] If there was a change in symptoms than a week, I could notice that (Participant 1, Groupe de discussion 2).

- 37 A contrario, un autre participant est plus critique et suggère que les graphiques pourraient être améliorés :

Maybe the chart could be improved because when you look at the long period [...] they're a bit of a mess to look at (Participant 3, Groupe de discussion 2).

- 38 Ces témoignages convergent vers une même idée : les données deviennent actionnables lorsqu'elles sont lisibles, contextualisées et liées à des objectifs personnels, favorisant ainsi l'autonomie et l'ajustement des comportements de santé. L'accès à des tendances encourage la mise en place de comportements créant une incitation à l'action, comme l'illustre ce témoignage à propos des informations liées à l'usage de la montre connectée :

It really helps to meet that goal every day and try to surpass it. And to avoid being too static (Participant 5, Groupe de discussion 2).

- 39 Dans ce cas, le suivi des données générées par la montre connectée crée une incitation à bouger, à atteindre son objectif quotidien. La visualisation des données agit ici comme un levier motivationnel. La régularité imposée par la mesure devient un vecteur d'autodiscipline, traduisant l'actionnabilité des données.

- 40 Les données ne se contentent pas d'informer, elles guident et structurent la routine du ou de la patient·e. Elles agissent comme des rappels et des conseils qui orientent les patient·es vers des pratiques d'autosoins. Cependant, si ces conseils manquent de personnalisation, ils risquent de rester génériques et de ne pas répondre précisément aux besoins individuels, ce qui limite leur pouvoir transformateur, comme l'illustre cet extrait d'entrevue :

Because I've had PD for 9, 10 years, I didn't find its symptom, symptom tracking to be useful for me because my symptoms haven't changed that much (Entrevue, Patient 6).

- 41 L'actionnabilité des données, c'est-à-dire leur capacité à se traduire en actions concrètes soutenant les pratiques d'autosoins, est une notion centrale dans le domaine des technologies numériques de santé qui se veulent personnalisées. Lupton (2016a et 2016b) montre que la quantification du corps peut fournir des repères objectifs qui permettent aux individus de prendre conscience de leurs comportements. Par exemple, lorsqu'un ou une patient·e constate une corrélation entre une durée de sommeil insuffisante et une aggravation de ses symptômes, ces données lui offrent une base pour adapter ses habitudes de vie. Ces usages qui sont partagés par les participant·es à l'étude soulignent que les données ne prennent sens que lorsqu'elles sont intégrées dans une logique d'action située, adaptée aux besoins du moment, aux préférences et aux contextes de vie des personnes. En effet, quand les données ne font pas sens, la surveillance génère du retrait :

I know how I feel in the morning, but it's interesting to see what the computer said [...] I didn't find that very helpful (Entrevue, Patient 18).

- 42 Ce participant exprime une ambivalence face au suivi quotidien ou aux graphiques, qui peuvent induire une forme de normalisation et laisser de côté son propre ressenti. D'autres ressentent une pression à atteindre des objectifs (nombre de pas, qualité du sommeil) ou évoquent une fatigue, une perte de spontanéité.

Les données comme moyen pour mieux communiquer avec son équipe de soins (connection)

- 43 Les données peuvent participer à créer des ponts entre le ou la patient·e, ses proches et les professionnel·les de santé. Certain·es patient·es expriment le souhait que les données puissent être partagées avec leur équipe de soins et ainsi servir de point d'appui pour discuter des symptômes ou des ajustements thérapeutiques :

Yeah, it can give me, if there's something that's not working out or whatever, it can give me a basis. [...] Maybe make a little note of them and bring them into your next appointment and give them to the nurse practitioner prior to seeing the physician, which is what I basically do anyways (Entrevue, Patient 2).

- 44 Ce souhait de partage souligne l'idée que les données ne sont pas uniquement destinées à l'autosuiti, mais aussi à l'enrichissement du dialogue avec les soignant·es. D'autres mettent en avant la connexion avec des technologies complémentaires, comme leur montre connectée, qui crée un écosystème de données intégrées :

J'apprécie vraiment le lien avec le Fitbit, je trouve que c'est de valeur, complémentaire. [...] c'est une intégration qui apporte (Entrevue, Patient 12).

- 45 L'intégration des technologies (montre, système de géolocalisation des services, etc.) crée un écosystème connecté qui renforce le sentiment d'appartenance à un réseau de soins. C'est ce que des patient·es mentionnent en faisant référence à l'outil de géolocalisation des services accessible par eCARE-PD 2.0 :

I was quite impressed with the connection to all different available services. [...] That was quite impressive. I did look at that and see what was immediately around me and I found that potentially to be very, very useful (Entrevue, Patient 10).
It's nice to know that maybe in Barrhaven or Stittsville, that there's someone who can do physio (Entrevue, Patient 6).

- 46 Les participant·es apprécient beaucoup la fonction de géolocalisation pour identifier des physiothérapeutes, centres communautaires ou autres. Cela renforce le sentiment d’être soutenu par un réseau de soins accessible. Cette valeur de connexion mise en avant par les patient·es souligne le fait que les données constituent un levier dans la coordination de soins. Cette fonction relationnelle des données — telle que l’expriment les patient·es dans leur récit — a le potentiel de favoriser une meilleure intégration et une meilleure coordination des soins. Dans les entrevues et groupes de discussion, les participant·es décrivent les données (ex. : entrées de journal, graphiques hebdomadaires ou mensuels, *care tips*) comme des supports pour dialoguer avec les soignant·es, préparer les rendez-vous ou discuter d’ajustements au traitement. Cela crée une valeur de connexion, où les données renforcent les liens entre patient·es, proches aidant·es et équipes de soins. Quand les données sont partagées, contextualisées et utilisées pour soutenir la communication avec l’équipe de soins, la surveillance soutient la relation de soins, comme on le voit ici.

Les données révélant des informations inattendues (discovery)

- 47 Les données permettent aux patient·es de découvrir des relations inattendues entre des symptômes et des éléments de leur quotidienneté ou de mieux comprendre leur propre corps qui est vécu comme imprévisible et parfois étranger. Cela est particulièrement pertinent dans le cas de la maladie de Parkinson, où les symptômes évoluent et interagissent de manière complexe avec d’autres facteurs (ex. : sommeil, alimentation, activité physique, stress). L’analyse des données permet parfois de mettre en lumière des éléments auxquels les patient·es n’avaient pas pensé et de découvrir des corrélations inattendues, comme l’influence de l’exercice régulier sur les symptômes ou la nécessité de modifier certains comportements pour améliorer leur sommeil ou leur mobilité :

By using the tracking, I realized I had pain (tightness) every day and did not realized before how much tightness (Entrevue, Patient 7).

- 48 La découverte de relations entre différentes données permet aux patient·es de mieux comprendre les dynamiques internes de leur corps. L’usage des données amène donc certain·es patient·es à développer une conscience corporelle accrue, car ces personnes commencent à relier certains comportements ou événements à des variations dans leurs symptômes. Cette prise de conscience peut se traduire de plusieurs manières, comme identifier des tendances ou confirmer des relations entre divers facteurs :

J’avais choisi la douleur pour en fait comprendre certaines choses par rapport à ma douleur. Comme j’écrivais, la douleur, c’est plus musculaire, c’est plus, je me sens plus tight, comme j’aurais dit en anglais. J’ai pu voir des spécificités sur certaines choses que je n’aurais pas pensé. Je ne suis pas une personne qui aime écrire des choses par rapport à moi. Ça m’a forcé à le faire, puis à comprendre aussi. Comme ma mobilité, il y a des jours que ça disparaît, des jours que ça n’allait pas bien. Ça m’a fait comprendre pourquoi ça n’est pas bien. Si j’ai beaucoup de douleurs, je ne peux pas mobiliser autant que si je n’en ai pas (Entrevue, Patiente 5).

- 49 Cependant, tous ne perçoivent pas cet effet de découverte de manière positive. Chez certaines personnes, ces découvertes peuvent augmenter l’anxiété, notamment en leur rappelant le caractère dégénératif de la maladie :

Also, I don’t find it helpful to track how I’m feeling about one topic or another, as eCARE-PD has no impact. If I could say “oh back in sept I was less happy with my weight but now I received these tips + they helped that would be ok”. But

unfortunately my history mostly consists of me being continually unhappy. How does it help me to track this? I just feel beaten down by my health problems. Going back and seeing my history of unhappiness merely reinforces it (Extrait du journal d'usage, Patient 4).

- 50 Ces témoignages révèlent une double dimension dans l'usage des données. D'une part, celles-ci semblent favoriser une prise de conscience corporelle accrue, permettant aux patient·es de relier leurs comportements à l'évolution de leurs symptômes, et d'autre part, elles peuvent accentuer l'anxiété en rappelant la nature dégénérative de la maladie. Certains participants expriment que le suivi de leurs activités et l'analyse des tendances permettent d'obtenir des repères concrets par rapport à leur état de santé. En donnant aux patient·es la possibilité d'identifier des corrélations entre leurs actions et leurs ressentis, ces outils favorisent une prise de décision informée et encouragent les pratiques d'autosoins. Néanmoins, pour certain·es patient·es, l'exposition continue à des statistiques — parfois déprimantes — rappelle sans cesse la progression dégénérative de la maladie. On peut voir que cette dynamique de découverte ne se fait pas sans tension.

Surveillance distribuée comme pratique dialogique et encapacitante

- 51 La surveillance distribuée mise en évidence dans le projet eCARE-PD dépasse la simple mesure des symptômes ou changements corporels. Elle repose sur une redistribution de l'attention entre différents acteurs — humains (patient·es, proches aidant·es, soignant·es) et non humains (technologie, algorithmes, données) — et instaure une forme de vigilance partagée. Contrairement à une surveillance centrée sur le contrôle ou la conformité à des normes médicales, elle s'apparente à une pratique dialogique, où les données, les ressentis corporels et les dispositifs techniques coproduisent du sens autour de la maladie. Autrement dit, les données de santé sont intégrées dans un échange réflexif entre le vécu corporel du ou de la patient·e, les graphiques et visualisations proposés par eCARE-PD 2.0 ainsi que les retours des équipes de soins ou des aidant·es. Cette dynamique dialogique permet aux patient·es de confronter leurs ressentis aux données et d'en faire sens, le tout soutenant l'émergence d'une compréhension située et partagée de la maladie et de ses symptômes. Ce processus se manifeste par une coactivité entre le ou la patient·e et le dispositif de télésuivi : la technologie capte, visualise et archive les données, tandis que le ou la patient·e les interprète, les confronte à son vécu et ajuste ses pratiques d'autosoins. Ce va-et-vient entre le corps vécu et le corps mesuré ouvre un espace de réflexivité, où les écarts entre ressenti et données ne sont pas nécessairement perçus comme des contradictions, mais comme des occasions de reconfiguration du rapport au corps.
- 52 L'analyse des *data valences* permet ici de qualifier les conditions dans lesquelles les données deviennent signifiantes pour les patient·es. Les six *data valences* identifiées par Fiore-Gartland et Neff révèlent que les données ne sont pas neutres : elles sont chargées d'attentes, d'émotions et de valeurs situées. Dans le cas de la maladie de Parkinson, les témoignages montrent que les patient·es attribuent aux données des fonctions multiples : elles servent à valider un ressenti, à justifier une action, à faciliter le dialogue avec les soignant·es ou à découvrir des corrélations inattendues entre symptômes et comportements ou actions. Ainsi, la surveillance distribuée soutient l'encapacitation, dans la mesure où elle permet un travail de construction de sens autour de la maladie. Elle ne repose pas uniquement sur la technologie, mais sur une

écologie du soin, où les données circulent, sont interprétées, discutées et intégrées dans des pratiques quotidiennes. Elle permet aux patient·es de mieux comprendre leur corps, d'anticiper les fluctuations de leurs symptômes et de s'engager dans une gestion active de leur santé.

- 53 Cette sur-veillance distribuée, telle qu'elle émerge à travers l'usage de eCARE-PD 2.0, invite à repenser les technologies de santé non comme des outils de contrôle, de persuasion, mais comme des médiateurs d'un dialogue continu entre le corps, la technique et le soin. Cela souligne l'importance de concevoir des dispositifs de télésuivi qui soutiennent la réflexivité, la personnalisation et la reconnaissance du vécu subjectif, en intégrant les *data valences* comme grille d'analyse et comme levier de conception. Une conception centrée sur la reconnaissance du vécu, où les données ne sont pas des injonctions à agir, mais des ressources pour penser, discuter et négocier le soin. Par conséquent, l'intégration des *data valences* comme grille d'analyse et comme levier de conception permet de reconnaître la pluralité des significations que les données peuvent prendre selon les contextes, les trajectoires et les expériences sensibles des patient·es.

Une sur-veillance équipée et (dé)connectée dans le cadre projet RealWorld4Clinic

- 54 Le projet RealWorld4Clinic met en lumière une tension fondamentale entre une sur-veillance équipée sans soutien humain et une sur-veillance équipée et connectée à l'équipe de soins. En effet, l'analyse des entretiens fait apparaître deux groupes distincts. D'un côté, les patient·es équipés de prothèses cardiaques connectées, mais non accompagnés, se retrouvent dans une situation de rupture entre leurs ressentis et les données accessibles aux médecins, ce qui accentue le sentiment de solitude face à la maladie. De l'autre, les patient·es équipés d'un dispositif d'autosurveillance et inclus dans un programme de télésuivi, pour qui la connexion régulière avec une équipe soignante permet une meilleure appropriation des données et une action plus rapide sur les traitements, ont une moins grande angoisse liée à l'incertitude des symptômes.

Une sur-veillance équipée, mais déconnectée de l'équipe de soins

- 55 Pour les patient·es qui n'ont pas directement accès aux données de leur prothèse, dont l'analyse est réservée aux rythmologues (cardiologues spécialisés dans l'étude des anomalies du rythme cardiaque), tout échange sur ces données de santé se fait donc a posteriori, parfois six mois ou un an après que les anomalies détectées par le ou la rythmologue ont eu lieu. En conséquence, les patient·es ont de grandes difficultés à établir un lien entre leurs sensations physiques et leurs données de santé du fait de ce décalage temporel :

Des fois je me sens mal. Comme dit mon cardiologue, je ne me sens pas bien. Je ne peux pas expliquer ce que c'est, mais je ne me sens pas bien. Alors je le note et quand je me fais contrôler, je dis : à telle période-là, qu'est-ce qu'il y a eu ? Ah non, rien. Tout est très bien. Par contre deux mois plus tard, boom, il y a une rafale (Entrevue, Patient 3).

- 56 La solitude face à des données de santé qui ne correspondent pas aux sensations peut mettre les patient·es en grande difficulté quand ils n'ont pas immédiatement d'échange

avec un ou une soignant·e. Ces deux problèmes associés créent un sentiment de solitude face à des symptômes et nous permettent d'opérer deux distinctions : 1) la temporalité d'analyse décalée crée du doute et de l'inconfort et 2) des décalages de correspondance entre données cliniques et données subjectives, comme dans le cas de la maladie Parkinson, crée une perturbation.

- 57 Les données dans ce cas ne permettent pas au patient ou à la patiente de comprendre ses sensations. Un doute sur la « véracité » des données émerge et le système ne permet pas d'actionner une « connexion » avec le personnel de santé afin de le lever. Il faut souligner que les patient·es ont parfois besoin d'avoir accès à certaines données mesurées par leur prothèse immédiatement. C'est le cas notamment des patient·es porteur·euses de défibrillateurs qui doivent veiller à ce que leur rythme cardiaque ne dépasse jamais la valeur de seuil de leur appareil (auquel cas l'appareil risquerait de considérer le rythme cardiaque comme pathologique et de délivrer un choc électrique). Ceux qui souhaitent pratiquer une activité physique sont donc encouragés à s'équiper d'une montre connectée :

[Comment vous savez si vous avez forcé ?] Bah je ne le sais pas. [pause] Bah je le sais... Je le vois un peu. Quand euh... j'ai programmé la montre. Je sais que je ne dois pas dépasser 110 pulsations. Donc j'ai mis une alarme sur la montre. [À cause du défibrillateur ?] Oui. 110. Si je dépasse les 110, c'est un peu le... Les cardiologues m'ont dit, m'ont tous dit... 110. Je ne dois pas dépasser 110.82 c'est là... 82 pulsations c'est le... pour le travail, pour le renforcer. Mais arriver à 110 c'est trop fort. Il ne peut pas aller au-dessus. Enfin il peut aller au-dessus, mais le défibrillateur va s'enclencher (Entrevue, Patient 24).

- 58 Pour ce patient, qui a du mal à ressentir précisément la fréquence de son rythme cardiaque, la montre connectée est un outil indispensable, qui lui fournit des informations précises (« clarté ») et lui permet de savoir ce qu'il doit faire précisément (« actionnabilité »). De plus, ces patient·es n'ont presque aucun contact avec les équipes chargées de la télésurveillance, sauf en cas de dysfonctionnement (potentiellement mortel) de leur appareil et, tel que nous l'observons ci-après, un dispositif de télésurveillance ne peut se suffire à lui-même. Les équipes sont constituées d'infirmiers hospitaliers, qui assurent ce suivi en plus de leur travail quotidien, ce qui peut les rendre peu réactifs. Un des patients a rapporté sa surprise de ne pas être contacté par l'hôpital après le déclenchement de son défibrillateur ; une autre d'avoir été contactée plusieurs mois après sa greffe cardiaque (et donc le retrait de son défibrillateur). Cela rejoint les observations réalisées en Suède par Liljeroos et ses collègues (Liljeroos, Thylén et Strömberg, 2020) qui notent par ailleurs que ce travail de télésuivi est peu reconnu par les médecins (bien que celui-ci leur prenne beaucoup de temps, du fait de la difficulté technique à paramétrer et à suivre les alertes provenant de modèles différents et de la documentation à produire pour chaque événement). À travers l'analyse des affects « négatifs » (déception, peur, colère) se révèlent ainsi des valences « en creux » : ce que les patient·es attendent du système de santé et qu'ils n'ont pas réussi à obtenir. Des transcriptions de ce type sont présentées dans le tableau suivant. La dimension Découverte n'est pas apparue dans les données, car ce projet avait pour but d'étudier les besoins des future·s patient·es de l'application de télésuivi de l'insuffisance cardiaque.

Tableau 1. Transcriptions extraites des entretiens menés dans le cadre du projet RealWorld4Clinic

<i>Data valence</i>	Exemple de transcription
Clarté	Je ne demande pas à pouvoir régler mon défibrillateur, hein, pas du tout. Mais des fois je sens une suspension cardiaque et je me dis : « oh, j'aimerais bien savoir si j'ai fait une arythmie, une fibrillation... ou si c'est purement psychologique (Entrevue, Patient 2)
Actionnabilité	Moi j'avais besoin de ça pour être rassuré quelque part, mais ce qui me paraissait bizarre c'est quand il y avait des alertes c'était moi qui faisais la démarche d'aller à l'information plutôt que l'information de venir à moi par le biais des gens qui surveillaient. Mais après j'ai appris comment se faisait la surveillance... (Entrevue, Patient 7)
Connexion	D'ailleurs cette nuit, dans le questionnaire des symptômes que me pose la tablette, il y a « Cette nuit, avez-vous été plus essoufflée ? ». Et justement cette nuit c'est arrivé, et de savoir qu'on est suivi comme ça, si vous voulez, on est un peu rassuré quoi. Je trouve que ça rassure ce programme (Entrevue, Patient 23).
Transparence	[Au sujet d'une application de santé testée par la personne pour noter ses médicaments et rappels de rendez-vous médicaux] Malheureusement, je me suis dit : demain s'il m'arrive quelque chose, les gens n'iront pas fouiller dans le téléphone. Donc je me suis dit : finalement, est-ce qu'elle m'est utile ? Du coup, pas tellement. Donc la seule chose que je fais maintenant, c'est que toutes ces informations, je les ai mises... en fait j'ai une sorte de petit carnet dans mon sac posé directement en me disant qu'en fait les gens auront l'identité de la personne, quoique ce soit pour voir. [...] C'est par expérience que je dis ça. J'avais fait un malaise au travail, justement à cause du cœur. Et finalement comme j'étais inconsciente, ils savaient pas trop quoi dire aux pompiers, quels médicaments, quels trucs. Ils disaient juste : elle a un problème au cœur, on sait qu'elle a une machine (Entrevue, Patient 6).
Véracité	J'étais un très très grand sportif. Je faisais entre 8 et 12h de sport par semaine. Et donc j'ai toujours eu [une montre connectée] quoi. Au moins pour savoir où j'en étais au niveau de la respiration, les battements, tout ça. Ce que j'avais fait sur tous mes entraînements quoi. [...] C'est sur mon téléphone directement. Je regarde au moins une fois par jour, on va dire, depuis l'incident. Avant je m'en servais surtout au moment où je faisais les efforts physiques. Ça n'allait pas plus loin quoi. Là, maintenant, depuis l'incident... Enfin depuis [date], avec l'infarctus un peu plus prononcé par rapport au premier, et le deuxième... enfin le troisième maintenant, je mélange toujours... là c'est un peu plus régulier... Je veux dire... Là j'en ai besoin quoi. Pour me rassurer, j'en ai besoin. Je vérifie la pression artérielle, les battements, et l'oxygénation. Généralement, c'est le matin quand je me lève, je regarde mon téléphone comme ça je sais la nuit comment ça s'est passé, et... le soir en fin de journée (Entrevue, Patient 24).

Une surveillance équipée et connectée à l'équipe de soins

- 59 Pour les patiente·s qui sont suivis par une équipe de soins d'un hôpital et qui font partie d'un programme de suivi à domicile supervisé par des infirmières spécialisées

dans le suivi de l'insuffisance cardiaque (nous avons pu interroger l'une d'entre elles), la situation est différente. Elles surveillent quotidiennement les données transmises par les patient·es par l'intermédiaire de leurs appareils connectés (le suivi des prothèses cardiaques en revanche n'est pas de leur ressort). Les patient·es sont contacté·es s par téléphone, non seulement en cas de signe d'aggravation de la maladie ou en l'absence de transmission, mais aussi pour maintenir le contact avec la clinique. Ces appels leur donnent l'occasion de discuter avec des professionnel·les de santé et ainsi de mieux comprendre le lien entre leurs données de santé et leurs sensations :

Je peux voir tous les résultats. Je vois aussi les courbes. Et d'ailleurs, dès qu'il y avait quelque chose qui n'allait pas... au début, des fois, tout d'un coup j'avais une perte de poids incroyable. Dans l'heure qui suivait, j'avais un coup de fil hein... des infirmières. Ah oui oui. À un moment donné, j'ai eu une baisse de tension, j'ai pas bien compris pourquoi d'ailleurs... La perte de poids, j'avais oublié que j'avais arrêté les compléments alimentaires qu'on me donnait, vous savez, parce que je m'étais vraiment très très affaiblie. Et donc j'avais des protéines à prendre. Et donc quand j'ai terminé mes protéines, forcément, ça s'est traduit tout de suite dans les deux jours qui ont suivi par une perte de poids. Évidemment que j'avais pas fait le lien. Et quand l'infirmière m'a appelée, je me suis dit : « Qu'est-ce qui a changé, qu'est-ce qui a changé... ? J'ai fini les protéines. » « Ah bon ? » Ouf, c'était ça (Entrevue, Patient 17).

- 60 Les patient·es qui font partie de ce programme doivent également participer à trois séances d'accompagnement thérapeutique. Les sujets abordés incluent l'annonce, les règles hygiéno-diététiques, l'activité physique, le parcours de soins et les consignes de sécurité. La télésurveillance est prévue pour six mois, renouvelables.
- 61 Cette technologie accompagnée par des professionnel·les de santé peut donc permettre aux patient·es d'apprendre à prendre soin de leur santé, en s'appropriant des données claires, transparentes et actionnables. C'est ce modèle que nous avons recommandé de suivre dans le cas de RealWorld4Clinic, d'autant que l'implant initial était conçu plutôt sur le modèle des prothèses cardiaques, et non comme un moyen d'accompagnement thérapeutique. Les patient·es partenaires ont suggéré la mise en place d'indicateurs très simples du risque de décompensation cardiaque inspirés de l'univers narratif du Code de la route, les feux tricolores : vert signifie *rien à signaler*, jaune, *contacter votre médecin référent pour contrôler le traitement*, rouge, *contacter les urgences*.
- 62 En résumé, l'application du modèle des *data valences* dans le contexte du projet RealWorld4Clinic permet de mieux comprendre certaines nuances visant les liens entre clarté/transparence, connexion/actionnabilité et véracité/découverte. La clarté des données est essentielle pour éviter un sentiment d'exclusion. Les patient·es qui ne comprennent pas les alertes fournies par leur prothèse cardiaque expriment un besoin accru de médiation humaine pour donner du sens aux données. L'absence de rétroaction rapide des équipes soignantes fragilise la valeur actionnable des données, laissant les patient·es dans une posture d'attente plutôt que d'anticipation. Et finalement, les témoignages révèlent une tension entre les données médicales et le ressenti subjectif des patient·es. Lorsque les chiffres ne confirment pas un malaise ressenti, la confiance dans la donnée et dans le suivi médical peut être mise à mal.
- 63 Par conséquent, ce projet illustre les conditions de l'intégration d'une technologie dans un processus d'autosoins pour le ou la patient·e. En d'autres termes, pour qu'elle soit signifiante et bienveillante, la sur-veillance exige une articulation claire entre les données produites par les dispositifs, leur interprétation par le ou la patient·e et le

soutien d'un ou d'une professionnel·le de santé. Il ne s'agit pas simplement de fournir des technologies, mais de construire un écosystème de suivi qui intègre les patient·es dans la boucle décisionnelle, en leur donnant les moyens de comprendre, d'agir et d'échanger avec leurs soignant·es.

Conclusion

- 64 Cet article visait à analyser dans quelle mesure les dispositifs de télésuivi peuvent favoriser un processus d'autosoins, en identifiant les conditions nécessaires à la production de sens autour des données de santé. En mobilisant le modèle de *data valences*, nous avons montré que ces données sont porteuses de significations situées, qui prennent forme dans les interactions entre patient·es, technologies et professionnel·les de santé. L'analyse croisée de deux projets — eCARE-PD (technologie d'autosoins pour les personnes atteintes de la maladie de Parkinson) et RealWorld4Clinic (implant de *monitoring* pour l'insuffisance cardiaque) — met en évidence que les dispositifs de télésuivi ne produisent pas nécessairement une forme unique de surveillance, mais bien une pluralité de modes de sur-veillance, selon la configuration du dispositif et les relations qu'il permet d'entretenir avec les soignant·es.
- 65 Dans le cas d'eCARE-PD, la sur-veillance apparaît comme distribuée et repose sur une coactivité entre le ou la patient·e, la technologie et l'écosystème de soins. Les patient·es s'engagent dans une activité réflexive qui consiste à interpréter leurs données, à les confronter à leurs ressentis et à ajuster leurs pratiques d'autosoins. Ce processus témoigne d'une reconfiguration du rapport au corps : les données deviennent médiatrices et soutiennent un travail d'interprétation, de construction de sens et de projection, participant ainsi à une vigilance partagée entre humains et technologie.
- 66 Dans RealWorld4Clinic, en revanche, la sur-veillance se révèle équipée, mais inégalement connectée à l'écosystème de soins : lorsque la technologie fonctionne sans médiation humaine immédiate, les patient·es expriment des sentiments d'incertitude ou de perte de sens face aux données. À l'inverse, lorsque le dispositif est intégré dans un programme de suivi accompagné par des professionnel·les de santé, les données deviennent claires, transparentes et actionnables, réduisant l'anxiété et favorisant une meilleure appropriation des informations. Ce contraste souligne que la valeur des données dépend étroitement des conditions relationnelles et organisationnelles dans lesquelles elles sont produites et circulent. Les analyses des deux projets illustrent le fait que les technologies qui vise l'autosoin se doivent d'introduire une veille active sur soi qui est à la fois une mise en visibilité du corps et des symptômes, une transformation du ressenti en données exploitables et une incitation à l'action.
- 67 Nos analyses permettent de dégager trois modalités fondamentales de la sur-veillance, qui structurent le rapport des patient·es aux données de santé dans le cadre des technologies d'autosoins. Ces modalités — représentation, objectivation et capacitation — constituent des dynamiques complémentaires par lesquelles les données prennent sens et soutiennent l'action.
- a. *La (re-)présentation* : L'un des rôles des technologies d'autosoins est de donner à voir l'évolution des symptômes, de matérialiser une expérience corporelle souvent diffuse. La représentation des symptômes à travers des courbes, des chiffres ou des historiques permet aux patient·es de mieux comprendre les fluctuations de leur état et d'y apporter du sens.

Cela leur donne accès à une mémoire de la maladie, un outil de comparaison dans le temps qui nourrit un travail réflexif. Cependant, cette visibilité impose aussi un cadre de perception normé, où seuls certains aspects du corps sont capturés et mesurés, au risque d'invisibiliser d'autres dimensions plus subtiles du vécu de la maladie.

b. *L'objectivation* : Elle consiste à matérialiser une sensation en un indicateur quantifié. Par exemple, la fatigue ou la rigidité musculaire, ressenties de manière subjective, peuvent être traduites en scores ou en tendances à partir de l'autosurveillance. Cette conversion du ressenti en données mesurables permet de structurer l'expérience de la maladie et de la rendre communicable aux soignant·es. Toutefois, elle peut aussi introduire une tension lorsque les chiffres ne correspondent pas à l'expérience vécue, le ou la patient·e est confronté à un écart entre son ressenti et la donnée objective, pouvant générer du doute.

c. *La capacitation* : La sur-veillance ne se limite pas à une observation passive, elle engage une capacité à agir sur son état de santé. L'accès aux données permet aux patient·es d'ajuster leurs comportements, de modifier leurs traitements ou d'anticiper des épisodes aigus. Dans le cas d'eCARE-PD 2.0, les notifications et conseils personnalisés incitent à une meilleure gestion des symptômes. Dans RealWorld4Clinic, la transmission des données aux soignant·es permet une prise en charge plus réactive. Toutefois, cette encapacitation est conditionnée par plusieurs facteurs : la lisibilité et l'accessibilité des données (comprendre les tendances, identifier des corrélations utiles), la possibilité d'agir concrètement sur son traitement ou son mode de vie et la prise en compte du contexte émotionnel du ou de la patient·e, pour éviter une surcharge cognitive ou anxiogène face à l'information.

68 Ces trois modalités permettent de qualifier la sur-veillance comme un processus à la fois technique, cognitif et relationnel, où l'autosoin se construit dans la circulation et la confrontation des données, des ressentis et des interactions humaines. Par conséquent, nous proposons de comprendre la sur-veillance comme redistribuant les rôles et les responsabilités entre le ou la patient·e, la technologie et les acteurs du soin. Loin d'un modèle de surveillance reposant sur le contrôle et la mesure, elle ouvre un espace dialogique où les données deviennent des médiateurs d'un apprentissage, c'est-à-dire apprendre à vivre avec un corps faillible et imprévisible.

69 L'usage du modèle des *data valences* s'est révélé particulièrement fécond pour évaluer la valeur encapacitante des dispositifs de télésuivi. En permettant de qualifier finement les attentes, les émotions et les tensions associées aux données de santé, ce cadre analytique offre une lecture située et nuancée du rapport que les patient·es entretiennent avec les données générées par des dispositifs numériques et algorithmiques. L'une des contributions de cet article réside dans l'articulation entre cette grille d'analyse et les pratiques d'usage réelles ou anticipées, révélant comment les données peuvent soutenir — ou au contraire entraver — l'autosoin.

70 Au-delà de l'évaluation des usages, le modèle des *data valences* constitue un levier pour la co-conception de technologies d'autosoins plus sensibles aux réalités vécues. En identifiant les valeurs prioritaires pour les patient·es — telles que la preuve, l'actionnabilité, la connexion ou la découverte —, il devient possible d'orienter la conception d'interfaces et de fonctionnalités qui ne cherchent pas à prescrire des comportements ou à induire des actions par des mécanismes de persuasion implicite, mais qui soutiennent un véritable travail de sens. Par conséquent, les dispositifs de télésuivi pourront appuyer l'autosoin lorsqu'ils contribuent à soutenir un double mouvement : d'une part, la représentation et l'objectivation du corps malade, permettant une meilleure compréhension de soi, d'autre part, l'encapacitation, entendue comme la possibilité d'agir, de dialoguer et de co-construire le soin. La sur-

veillance, en ce sens, ne relève pas d'un simple transfert de responsabilité vers l'individu, mais d'une forme de vigilance partagée, à la fois humaine et algorithmique, où le travail sur et avec la donnée crée un espace de rencontre entre le corps et ses symptômes, la technique et la relation de soins.

BIBLIOGRAPHIE

ANDRIEU Bernard (2016), « L'osmose technique avec son corps viv@nt : une auto-santé connectée du patient immersif », *L'Évolution psychiatrique*, 81, p. 253-265.

AKRICH Madeleine (1991), « De la sociologie des techniques à une sociologie des usages », *Techniques et culture*, 16, p. 31-54.

BARDOT Sandra, Bradley REY, Lucas AUDETTE, Kevin FAN, Da-Yuan HUANG, Jun LI, Wei LI et Pourang IRANI (2022), « One ring to rule them all: An empirical understanding of day-to-day smartring usage through in-situ diary study », *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3), p. 1-20.

BOL Nadine, Eline Suzanne SMIT et Mia Liza A. LUSTRIA (2020), « Tailored health communication: Opportunities and challenges in the digital era », *Digital Health*, 6, DOI : <https://doi.org/10.1177/2055207620958913>.

BRAYET Élise (2022), « Santé (dé)connectée ? Pistes de réflexions et expériences de l'altérité en contexte de santé connectée », *Interfaces numériques*, 11(3), p. 1-24.

BRAYET Élise et Isabelle QUEVAL (2023), « Corps écrans, corps en chiffres : se perdre ou se trouver. Quels enjeux pour l'éducation à la santé ? », dans Renaud HÉRITIER et Éric MUTABAZI (dir.), *La relation à autrui en régime numérique : sentir, apprendre et exister*, Lyon, Chronique Sociale, p. 51-67.

BERSON Josh (2015), *Computable bodies: Instrumented life and the human somatic niche*, Londres, Bloomsbury.

BRAUN Virginia et Victoria CLARKE (2024), « Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: A review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG) », *Palliative Medicine*, 38(6), p. 608-616.

BRAUN Virginia, Victoria CLARKE, Nicola HAYFIELD, Lisa DAVEY et Emily JENKINSON (2023), « Doing reflexive thematic analysis », dans Sofie BAGER-CHARLESON et ALISTAIR MCBWEATH (dir.), *Supporting Research in Counselling and Psychotherapy: Qualitative, Quantitative, and Mixed Methods Research*, Cham, Springer International Publishing, p. 19-38.

BRAUN Virginia et Victoria CLARKE (2006), « Using thematic analysis in psychology », *Qualitative Research in Psychology*, 3(2), p. 77-101.

CAHOUR Betty et Christian LICOPPE (2010), « Confrontations aux traces de son activité. Compréhension, développement et régulation de l'agir dans un monde de plus en plus réflexif », *Revue d'anthropologie des connaissances*, 4(4-2), DOI : <https://doi.org/10.3917/rac.010.0243>.

- CANGUILHEM Georges (2013), *Le normal et le pathologique*. 12^e édition, Paris, Presses universitaires de France, coll. « Quadrige ».
- CAMPBELL Karen A., Emily ORR, Patricia DUREPOS, Linda NGUYEN, Lucy LI, Charles WHITMORE, Peter GEHRKE, Laura GRAHAM et Shannon M. JACK (2021), « Reflexive thematic analysis for applied qualitative health research », *The Qualitative Report*, 26(6), p. 2011-2028, DOI : <https://doi.org/10.46743/2160-3715/2021.5010>.
- CHENEY-LIPPOLD John (2017), *We are data: Algorithms and the making of our digital selves*, New York, New York University Press, DOI : <https://doi.org/10.18574/nyu/9781479888702.001.0001>.
- DAGIRAL Éric (2019), « Extension chiffrée du domaine du perfectionnement ? La place des technologies de quantification du soi dans les projets d'auto-optimisation des individus », *Ethnologie française*, 49(4), p. 719-734.
- DAGIRAL Éric, Bertrand DERBEZ, Abhijit PEERBAYE et Dominique SAINT-MARC (2022), « Entre technologies de santé et soin des technologies : produire des matérialités soignantes », *Anthropologie & santé*, 25, DOI : <https://doi.org/10.4000/anthropologiesante.12308>.
- DAVAT Ambre et Fabienne MARTIN-JUCHAT (2023), « Enjeux info-communicationnels du co-design des technologies de santé », *Communication et organisation*, 63, DOI : <https://doi.org/10.4000/communicationorganisation.11877>.
- DAVAT Ambre, Fabienne MARTIN-JUCHAT et Thierry MENISSIER (2024), « Co-design with affect stories and applied ethics for health technologies », *Frontiers in Communication*, 9, DOI : <https://doi.org/10.3389/fcomm.2024.1327711>.
- DÉLÉTROZ Carole, Pamela BOU-MALHAB, Patrick BODENMANN et Marie-Pierre GAGNON (2022), « Chapitre 1.6. Les spécificités de la littératie en santé numérique des patients à l'heure d'Internet et du numérique », dans Patrick BODENMANN, Yves JACKSON, Francis VU et Hans WOLFF, *Vulnérabilités, diversités et équité en santé*, Chêne-Bourg, Médecine & Hygiène, p. 75-82, DOI : <https://doi.org/10.3917/mh.boden.2022.01.0075>.
- DESNOS Michel et Pierre JOURDAIN (2020), « Télémédecine : une solution d'avenir pour l'insuffisance cardiaque ? », *Bulletin de l'Académie nationale de médecine*, 204(8), p. 817-825, DOI : <https://doi.org/10.1016/j.banm.2020.07.034>.
- DUMAS Anne et Fabienne MARTIN-JUCHAT (2022), *Les communications affectives en organisations*, Bruxelles, De Boeck.
- DEWEY John (2011), *La formation des valeurs*, Paris, La Découverte, coll. « Les Empêcheurs de penser en rond ».
- FIORE-SILFVAST Brittany et Gina NEFF (2013), « What we talk about when we talk data: Valences and the social performance of multiple metrics in digital health », *Ethnographic Praxis in Industry Conference Proceedings*, 1, p. 74-87.
- FIORE-GARTLAND Brittany et Gina NEFF (2015), « Communication, mediation, and the expectations of data: Data valences across health and wellness communities », *International Journal of Communication*, 9, p. 19.
- GEERTZ Clifford (1983), *Local knowledge: Further essays in interpretive anthropology*, New York, Basic Books.
- GROSJEAN Michèle et Michèle LACOSTE (2010), *Communication et intelligence collective. Le travail à l'hôpital*, Paris, Presses universitaires de France.

GROSJEAN Sylvie (2024), « Susciter des récits d'usage de dispositifs d'autosoins "intelligents" par la méthode du récit à achever », *Réseaux*, 243(1), p. 143-182.

GROSJEAN Sylvie, Lauriane GIGUÈRE, Jason LEACH et Tiago MESTRE (2025), « Empowering self-care: leveraging user insights to co-design AI-powered self-care technology for Parkinson's disease », *FACETS*, 10(1-18), DOI : <https://doi.org/10.1139/facets-2024-0324>.

GROSJEAN Sylvie, Jean-Luc CIOCCA, Amélie GAUTHIER-BEAUPRÉ, Emely POITRAS, David GRIMES et Tiago MESTRE (2022), « Co-designing a digital companion with people living with Parkinson's to support self-care in a personalized way: The eCARE-PD Study », *Digital Health*, 8, DOI : <https://doi.org/10.1177/20552076221081695>.

HÄIKIÖ Juha, Sari YLI-KAUHALUOMA, Minna PIKKARAINEN, Marika LIVARI et Timo KOIVUMÄKI (2020), « Expectations to data: Perspectives of service providers and users of future health and wellness services », *Health and Technology*, 10, p. 621-636.

JARRAHI Mohammad H., Christine GOAY, Samuel ZIRKER et Yi ZHANG (2021), « Using digital diaries as a research method for capturing practices in situ », dans Gillian SYMON Gillian, Katrina PRICHARD et Christine HINE (dir.), *Research methods for digital work and organization: Investigating distributed, multi-modal, and mobile work*, Oxford, Oxford University Press.

JARKE Juliane et Sarah BÜCHNER (2024), « Who cares about data? Data care arrangements in everyday organizational practice », *Information, Communication & Society*, 27(4), p. 702-718.

JOKELA Tero, Jukka OJALA et Thomas OLSSON (2015), « A diary study on combining multiple information devices in everyday activities and tasks », dans *Proceedings of the 33rd Annual ACM Conference on human factors in computing systems*, New York, Association for Computing Machinery (ACM), p. 3903-3912, DOI : <https://doi.org/10.1145/2702123.2702211>.

KREPS Gary L. et Linda NEUHAUSER (2013), « Artificial intelligence and immediacy: designing health communication to personally engage consumers and providers », *Patient Education and Counseling*, 92(2), p. 205-210.

LAUDE Anne et Claude RAMBAUD (dir.) (2013), « La révolution du numérique et l'éducation thérapeutique du patient », *Journal du droit de la santé et de l'assurance-maladie (JDSAM)*, 4(4), p. 13-16, DOI : <https://doi.org/10.3917/jdsam.hs1.0014>.

LICOPPE Christian, Lorenzo DRAETTA et Anne DELANOË (2013), « Des "smart grids" au "quantified self". Technologies réflexives et gouvernement par les traces, une étude de cas sur la consommation électrique en milieu domestique », *Intellectica. Revue de l'Association pour la recherche cognitive*, 59(1), p. 267-290.

LILJEROOS Mona, Ingrid THYLÉN et Anna STRÖMBERG (2020), « Patients' and nurses' experiences and perceptions of remote monitoring of implantable cardiac defibrillators in heart failure: Cross-Sectional, descriptive, mixed methods study », *Journal of Medical Internet Research*, 22(9), DOI : <https://doi.org/10.2196/19550>.

LOMBORG Stine, Henriette LANGSTRUP et Thomas-Olof ANDERSEN (2020), « Interpretation as luxury: Heart patients living with data doubt, hope, and anxiety », *Big Data & Society*, 7(1), DOI : <https://doi.org/10.1177/205395172092443>.

LUPTON Deborah (2016a), *The quantified self: A sociology of self-tracking*, Cambridge, Polity.

LUPTON Deborah (2016b), « Digital health technologies and digital data: New ways of monitoring, measuring and commodifying human bodies », dans *Research handbook on digital transformations*, Cheltenham, Edward Elgar Publishing, p. 85-102.

- LUPTON Deborah (2017), « Feeling your data: Touch and making sense of personal digital data », *New Media & Society*, 19(10), p. 1599-1614.
- MATHIEU-FRITZ Alexandre et Caroline GUILLOT (2017), « Les dispositifs d'autosurveillance du diabète et les transformations du "travail du patient" », *Revue d'anthropologie des connaissances*, 114(4), p. 641-675, DOI : <https://doi.org/10.3917/rac.037.0641>.
- MATHIEU-FRITZ Alexandre et Nicolas GÉRARD (2023), « Les reconfigurations du "travail du patient" et de la relation thérapeutique lors de l'intégration d'un dispositif de télésurveillance médicale. Le cas de la diabétologie », *Sciences sociales et santé*, 41(2), p. 75-100.
- MAYÈRE Anne (2018), « Patients projetés et patients en pratique dans un dispositif de suivi à distance », *Réseaux*, 207(1), p. 197-225.
- MOL Annemarie (2008), *The logic of care. Health and the problem of patient choice*, Londres, Routledge.
- MOL Annemarie (2002), *The body multiple : Ontology in medical practice*, Durham, Duke University Press.
- NEFF Gina, Anissa TANWEER, Brittany FIORE-GARTLAND et Laura OSBURN (2017), « Critique and contribute: A practice-based framework for improving critical data studies and data science », *Big Data*, 5(2), p. 85-97.
- NIKOU Shahrokh, Wendy AGAHARI, Wilma KEIJZER-BROERS et Mark DE REUVER (2020), « Digital healthcare technology adoption by elderly people: A capability approach model », *Telematics and Informatics*, 53, DOI : <https://doi.org/10.1016/j.tele.2019.101315>.
- NUNES Francisco (2017), *The everyday life with Parkinson's and self-care technologies*, Dissertation, Technische Universität Wien, DOI : <https://doi.org/10.34726/hss.2017.47508>.
- ONNO Julien (2022), « L'amélioration de soi en régime numérique : les ambiguïtés du *Quantified Self*, entre discours et pratiques », *Socio-anthropologie*, DOI : <https://doi.org/10.4000/socio-anthropologie.12712>.
- OUDSHOORN Nelly (2008), « Diagnosis at a distance: The invisible work of patients and healthcare professionals in cardiac telemonitoring technology », *Sociology of Health & Illness*, 30(2), p. 272-288, DOI : <https://doi.org/10.1111/j.1467-9566.2007.01032.x>.
- PIERRON Jean-Philippe (2020), « Le soin dans la technique ou ce que les nouvelles technologies numériques font au soin », *Ethics, Medicine and Public Health*, 15, DOI : <https://doi.org/10.1016/j.jemep.2020.100597>.
- PINK Sarah, Debora LANZENI et Heather HORST (2018), « Data anxieties: Finding trust in everyday digital mess », *Big Data & Society*, 5(1), DOI : <https://doi.org/10.1177/2053951718756685>.
- POLANYI Michael (1966), « The logic of tacit inference », *Philosophy*, 41(155), p. 1-18.
- PRAINSACK Barbara (2017), *Personalized medicine: Empowered patients in the 21st Century?*, New York, New York University Press, DOI : <https://doi.org/10.18574/nyu/9781479814879.001.0001>.
- PRÉVÔT-HUILLE Hélène (2018), « Évaluation d'usage : méthodes et perspectives », dans Hervé MICHEL, Hélène PREVOT-HUILLE et Robert PICARD (dir.), *L'avenir des Silver Tech : conception, usage et évaluation*, Paris, Presses de l'EHESP, p. 241-265.
- RUCKENSTEIN Minna et Natasha D. SCHÜLL (2017), « The datafication of health », *Annual Review of Anthropology*, 46(1), p. 261-278.

- RUCKENSTEIN Minna (2014), « Visualized and interacted Life: Personal analytics and engagements with data doubles », *Societies*, 4(1), p. 68-84, DOI : <https://doi.org/10.3390/soc4010068>.
- SAILLANT Francine et Éric GAGNON (1996), « Le self-care : de l'autonomie-libération à la gestion du soi », *Sciences sociales et santé*, 14(3), p. 17-46.
- SMITH Gavin J. D. et Barbara VONTETHOFF (2016), « Health by numbers? Exploring the practice and experience of datafied health », *Health Sociology Review*, 26(1), p. 6-21, DOI : <https://doi.org/10.1080/14461242.2016.1196600>.
- VENKATESH Viswanath, Michael G. MORRIS, Gordon B. DAVIS et Fred D. DAVIS (2003), « User acceptance of information technology: Toward a unified view », *MIS quarterly*, 27(3), p. 425-478.
- VERGHESE Abraham (2008), « Culture shock — patient as icon, icon as patient », *The New England Journal of Medicine*, 359(26), p. 2748-2751.

NOTES

1. Dans cet article, nous ne nous focaliserons pas sur les types de données produites par les deux dispositifs. Ces données visualisables sont dépendantes des choix ergonomiques et de design des interfaces de télésuivi, questionnements précédemment publiés pour les autrices (Davat et Martin-Juchat, 2023 ; Davat, Martin-Juchat et Ménissier, 2024 ; Grosjean *et al.*, 2022 ; Grosjean *et al.*, 2025). Ce qui nous intéresse dans le cadre de cet article, c'est de mettre en évidence l'importance de comprendre, puis de catégoriser les attentes de valeurs ajoutées des données, à partir du recueil des vécus des patient·es.
2. Le projet eCARE-PD a été lancé par les docteurs Mestre et Grimes (neurologues) et a été financé par les Instituts de recherche en santé du Canada (IRSC) et Parkinson Canada. Dans le cadre de cet article, nous porterons notre attention sur la version 2.0 de eCARE-PD, qui est le résultat d'un processus de co-conception amorcé en 2019. Le projet se poursuit jusqu'en 2029, avec le codéveloppement d'un système de recommandation conversationnel (*LLM-based CRS specialized in Parkinson*).
3. RealWorld4Clinic est un consortium européen de 2020 à 2022 associant une quinzaine de partenaires publics et privés. Son but était d'accompagner le développement de MyHeartSentinel, un implant de *monitoring* de la santé cardiaque connecté issu de recherches menées au sein du laboratoire TIMC de l'Université Grenoble Alpes et qui devait être commercialisé par la *start-up* SentinHealth.
4. Notre traduction. Texte original : « the multi-dimensional expectations for and values around data operating within different data ecologies ».
5. L'European Institute of Innovation and Technology est la branche santé de l'Institut européen d'innovation et de technologies.
6. Nous ne présenterons pas de capture écran des visualisations de données pour ce projet pour des raisons de confidentialité et de droits d'auteur.
7. Précisons qu'un des *pacemakers* ne télétransmettait pas de données et que trois des patient·es avaient reçu une greffe cardiaque au moment de l'entretien et ne portaient donc plus de prothèse cardiaque.
8. Madopa est un *Living Lab* de l'EIT Health depuis décembre 2014 et partenaire associé depuis 2016 : <https://www.madopa.fr/contact/>.
9. Tous les extraits tirés du projet eCARE-PD seront dans la langue des participant·es, soit en français ou en anglais. Nous avons laissé les extraits dans la langue des participant·es afin de garder leurs expressions et de rester au plus près de leurs témoignages. Nous avons donc pris le

parti dans cet article de ne pas les traduire et de les laisser dans leur langue d'origine (alternant témoignages en français et témoignages en anglais).

RÉSUMÉS

Cet article explore comment les données de santé générées par des technologies algorithmiques de télésuivi peuvent soutenir un processus d'autosoins chez les patient·es atteints de maladies chroniques. En mobilisant le modèle des *data valences* (Fiore-Silfvast et Neff, 2013 ; Fiore-Gartland et Neff, 2015), nous analysons les conditions dans lesquelles ces données acquièrent une valeur située, en fonction des attentes, des émotions et des usages concrets des patient·es. L'étude s'appuie sur deux projets de co-conception de dispositifs de télésuivi : *eCARE-PD* (maladie de Parkinson) et *RealWorld4Clinic* (insuffisance cardiaque). L'analyse met en évidence le fait que les dispositifs de télésuivi ne produisent pas une forme unique de surveillance, mais bien une pluralité de modes de « sur-veillance », selon les configurations techniques, relationnelles et organisationnelles dans lesquelles ils s'inscrivent. Cette sur-veillance ne relève pas d'un simple contrôle, mais s'inscrit dans une dynamique distribuée et dialogique, où patient·es, technologies et professionnel·le·s de santé co-construisent une vigilance partagée, porteuse d'encapacitation.

This article explores how health data generated by algorithmic remote monitoring technologies can support a process of self-care in patients with chronic diseases. Drawing on the data valences model (Fiore-Silfvast and Neff, 2013; Fiore-Gartland and Neff, 2015), we analyze the conditions in which these data acquire specific value, based on patients' expectations, emotions, and concrete uses. The study is based on two co-design telemonitoring devices: *eCARE-PD* (Parkinson's disease) and *RealWorld4Clinic* (heart failure). Our analysis shows that remote monitoring devices do not produce a single form of monitoring, but rather a plurality of "over-monitoring" modes, according to the respective technical, relational, and organizational configurations. This over-monitoring is not simply a matter of control, but is part of a distributed and dialogical dynamic, in which patients, technologies, and healthcare professionals work together to create a method of shared monitoring, which fosters empowerment.

Este artículo presenta de qué manera los datos relativos a la salud, producidos por las tecnologías algorítmicas de monitoreo remoto pueden apoyar un proceso de autocuidado en los pacientes que padecen enfermedades crónicas. Tomando como base el modelo de *data valences* (Fiore-Silfvast y Neff, 2013; Fiore-Gartland y Neff, 2015), analizamos las condiciones en las cuales estas informaciones adquieren un valor localizado, en función de las expectativas, emociones y usos concretos de los pacientes. El estudio se apoya en dos proyectos de concepción en colaboración del dispositivo monitoreo remoto: *eCARE-PD* (enfermedad de Parkinson) y *RealWorld4Clinic* (insuficiencia cardíaca). El análisis pone en evidencia que los dispositivos de monitoreo remoto no producen una forma única de vigilancia, sino más bien una pluralidad de modos "vigilancia", según las configuraciones técnicas, de relaciones y organización en las que se enmarcan. Esta vigilancia no se limita a un simple control, sino que forma parte de una dinámica distribuida y de un diálogo, en el que el paciente, las tecnologías y profesionales del sector de la salud construyen en colaboración, una vigilancia compartida, portadora de empoderamiento.

INDEX

Mots-clés : datafication, télésuivi, autosoin, data valence, encapacitation

Palabras claves : autocuidado, dataficación, data valences, empoderamiento, monitoreo remoto

Keywords : self-care, datafication, data valence, empowerment, remote monitoring

AUTEURS

SYLVIE GROSJEAN

Sylvie Grosjean est professeure titulaire au Département de communication de l'Université d'Ottawa. Courriel : sylvie.grosjean@uottawa.ca

FABIENNE MARTIN-JUCHAT

Fabienne Martin-Juchat est professeure des universités en sciences de l'information et de la communication à l'Université Grenoble Alpes. Courriel : fabienne.martin-juchat@univ-grenoble-alpes.fr

AMBRE DAVAT

Ambre Davat est maîtresse de conférences en sciences de l'information et de la communication à l'École nationale supérieure des sciences de l'information et des bibliothèques. Courriel : ambre.davat@enssib.fr

Donnée administrative, ouvre-toi !

Le travail scientifique sur des données de subventions associatives

Cécile Rodrigues et Thomas Chevallier

- 1 Le mouvement d'émergence et de développement d'initiatives, de réglementations et de dispositifs visant à favoriser l'ouverture des données publiques (ou *open data*) date des années 2000 (Goëta et Mabi, 2014 ; Labelle et Le Corf, 2012)¹. En France, il a donné lieu à certaines politiques publiques, en particulier impulsées par la Loi pour une République numérique de 2016, qui instaure le principe d'ouverture des données publiques par défaut. Ces efforts sont pourtant loin d'être univoques tant ils servent des points de vue et intérêts variables, sinon contradictoires. Par ailleurs, l'application de ce principe d'ouverture des données publiques par défaut implique des moyens, notamment en termes de personnel pour assurer sa mise en œuvre concrète. L'informatisation généralisée de la société et le développement rapide des technologies de l'information et de la communication ont certes rendu possible la diffusion des données produites par le travail des administrations publiques au plus grand nombre. Reste à savoir comment cette ouverture des possibles est investie par différents publics, parmi lesquels les pouvoirs publics, les agent·es des collectivités, le secteur marchand, la société civile — y compris les sphères militantes et, nous concernant, le monde de la recherche.
- 2 Pour promouvoir l'*open data*, les discours s'appuient tantôt sur des objectifs de modernisation de l'action publique, tantôt sur l'idée de favoriser l'innovation économique, ou même encore de transparence et d'émancipation citoyenne (Baudot, Marrel et Nonjon, 2015 ; Paquienséguy et Dymytrova, 2018). Surtout, la prise en charge de l'enjeu par la politique publique aux différents échelons de gouvernement a eu pour conséquence « le passage d'une logique de demande à une logique d'offre » (Courmont, 2017, p. 19). Or, cette captation en amont n'a pas manqué d'avoir des effets sur la recherche puisque les travaux existants portent surtout sur l'offre publique d'*open data* en montrant que la réutilisation des données ouvertes reste souvent faible et en identifiant les obstacles à lever afin de favoriser leur appropriation par le plus grand nombre (parmi d'autres : Janssen, Charalabidis et Zuiderwijk, 2012 ; Martin, 2014 ; Conradie et Choenni, 2014). Mais la mise à disposition des données et leur réutilisation n'ont rien d'évident tant celles-ci sont marquées par les « infrastructures

informationnelles » dans lesquelles elles ont été produites (Edwards *et al.*, 2009 ; Bowker *et al.*, 2010).

- 3 En s'inscrivant dans les travaux portant sur les pratiques, les interactions et les processus occasionnés par la production, la diffusion et l'utilisation des données ouvertes (Courmont, 2015 ; Denis et Goëta, 2017 ; Ruijter *et al.*, 2020), le présent article propose de penser le travail en coulisses et les différentes étapes et épreuves à travers lesquelles des données peuvent être « détachées » de leur contexte institutionnel d'origine (Labelle et Le Corf, *op. cit.*) et utilisées par d'autres utilisateur·rice·s pour d'autres objectifs que ceux initialement pensés. Une telle perspective relationnelle sous-tend que les données publiques participent à façonner l'environnement qui les produit et les exploite (Courmont, 2017). Ainsi, considérées comme des construits sociaux, les données sont le produit d'un travail de production et de diffusion fourni par les agent·e·s dans les administrations en fonction de contraintes et de choix. Mais ce travail implique aussi un ensemble d'échanges et de transactions ayant des effets sur les rapports entretenus non seulement en interne des administrations, mais aussi avec des interlocuteur·rice·s diversifié·e·s, intéressé·e·s par le processus d'ouverture des données. Ce processus peut donc être interprété comme un rapport social, voire comme un rapport de force (Baudot, 2015).
- 4 À rebours d'une approche par l'offre et les politiques d'*open data* se centrant sur les enjeux de transparence, de démocratisation ou encore de valorisation économique des données, nous proposons d'aborder l'enjeu de l'ouverture des données sous l'angle des conditions de possibilité d'une appropriation des données par de potentiel·le·s usager·ère·s — ou comment les « données » (*data*) sont « saisies » (*capta*) (Becker, 1952, p. 278-279). D'une part, l'appropriation de données « métiers » implique de disposer des ressources, en termes d'expertise et de temps surtout. On pense alors à tout un monde privé des professionnel·le·s de la *data* qui se développe aujourd'hui, notamment autour de l'enjeu des *big data* et du *data mining* (Ollion et Hobeika, 2018) ou encore du monde militant des « statactivistes » (Bruno, Didier et Prévieux, 2014). Mais les usager·ère·s doivent, d'autre part, y voir un potentiel intérêt, qui peut correspondre à des visées économiques, scientifiques, militantes. Certaines organisations de la société civile telles que des associations, ONG, mouvements sociaux ou groupes de pression ont également, notamment dans le domaine environnemental, joué un rôle-clé dans l'ouverture de certaines données administratives (Domard, 2012 ; Goëta et Mabi, *op. cit.*). Mais *quid* du rôle qu'est susceptible de jouer la recherche publique dans de tels processus d'ouverture ? C'est à ce rôle, encore peu questionné, de la place de la recherche dans la production de données ouvertes que le présent article propose de s'intéresser à travers l'exemple spécifique de l'ouverture des données sur les subventions publiques d'une municipalité dédiées aux associations.

Encadré 1. Une enquête sur les libertés associatives

C'est à partir d'une recherche sur les relations entre les associations et les pouvoirs publics, dont l'un des axes concernait l'étude des financements publics des associations, que les enjeux et questions sur l'ouverture des données publiques et la place de la recherche dans ce processus se sont manifestés. L'un des volets de cette recherche portait sur une analyse statistique visant à rendre compte des transformations des modalités de financement des associations par les acteurs publics, notamment dans la continuité du travail entrepris par Julien Talpin et Pierre Bonnevalle à Roubaix (2023). En partant de la connaissance existante sur le

monde associatif et l'état de son partenariat avec les pouvoirs publics, nous avons identifié un manque concernant la question des financements. Si l'argent est systématiquement désigné comme une dimension-clé de la transformation des mondes associatifs, peu de travaux proposent de « mettre le nez dans les chiffres ». Et pour cause : c'est un manque de chiffres, notamment à un niveau localisé, qui a motivé la sollicitation d'une grande ville d'un peu plus de 200 000 habitant·es du nord de la France dans l'optique d'« ouvrir » de telles données.

Le choix de cette collectivité a été commandé par la connaissance du territoire, des spécificités du paysage associatif et de l'action publique sur laquelle nous pouvions nous appuyer. Cet échelon territorial, le plus proche des habitant·es, implique un dialogue plus régulier avec les associations, notamment celles qui revendiquent un ancrage dans le territoire et donc une capacité à dire ce qui est bon pour lui. Par ailleurs, cette municipalité est gouvernée par le même parti depuis environ un siècle. Dans la région, ce parti a largement construit son hégémonie, aux différentes étapes de son développement, sur la construction d'un lien étroit avec la société civile. Ce lien a pu prendre la forme de relations clientélares, bien que plusieurs tendances montrent qu'une nouvelle génération de responsables qui s'est imposée dans les années 2000 a cherché à s'inscrire en faux face à ces tendances et à moderniser les rapports dans le sens du partenariat et de la professionnalisation.

- 5 Démarrée en 2021, cette recherche a bénéficié de l'appui d'un financement « recherche participative » de l'Université. Si une approche « participative » en bonne et due forme n'a pas été mise en œuvre, le projet a reposé sur la construction d'un dialogue, aux différentes étapes du processus de recherche, avec des têtes de réseaux associatives régionales ainsi qu'avec des responsables de la municipalité étudiée. Les différents moments de partage des résultats ont permis de prendre la mesure de l'enjeu potentiel que pouvait représenter l'ouverture des données sur les subventions ou, autrement dit, de la demande sociale qui l'accompagnait.
- 6 Dans cet article, nous montrons comment, en cherchant à nourrir le débat scientifique et public sur le gouvernement et l'autonomie des associations (voir encadré 1), notre démarche d'enquête sociologique sur les subventions municipales dédiées aux associations a permis de souligner l'intérêt d'une approche d'ouverture des données « par la demande », c'est-à-dire conduite à travers une coopération entre les producteur·rice·s et les utilisateur·rice·s. Une telle approche conduit à renverser la perspective qui domine les politiques de l'*open data* en montrant que la « déspécification » des données administratives conduit à leur appauvrissement et qu'elle est finalement un horizon vain tant l'ouverture des données ne peut pas se faire sans ses utilisateur·rice·s. Notre hypothèse est donc que c'est plutôt par une démarche de contextualisation que la recherche publique est susceptible de jouer un rôle d'intermédiaire décisif dans l'ouverture des données publiques sur les subventions dédiées aux associations.
- 7 Dans un premier temps, nous exposerons les conditions politiques, administratives et techniques qui ont permis l'accès à ces données. Après une phase de négociation et grâce à l'aval de l'élue adjointe à la vie associative, ces données nous ont été fournies par la personne responsable du contrôle de gestion des subventions. Nous aborderons ensuite les opérations de nettoyage et les échanges avec les agent·es qui ont été

nécessaires afin de rendre ces données « métiers » exploitables dans le cadre de nos questionnements scientifiques. À travers ces opérations, l'ouverture des données était indissociable d'une démarche de contextualisation de ces dernières. Enfin, nous présenterons les voies à partir desquelles nous avons entrepris d'enrichir ces données afin d'en étendre la portée. Ce faisant, nous montrerons que notre travail de recherche, bien que motivé par des intérêts particuliers, permet de formuler des pistes de réflexion quant aux manières de produire une politique d'*open data* orientée vers la demande, a fortiori concernant cette thématique des subventions publiques dédiées aux associations.

Des données obtenues : les conditions d'accès aux données de gestion produites par une institution

- 8 Depuis leurs débuts, les sciences sociales se sont nourries de données administratives qu'elles ont appris à comprendre et à manier en vue de les mettre au service de questionnements propres. Pour ce faire, elles ont développé un ensemble d'opérations, de méthodes et de savoir-faire constituant une part importante, bien que peu valorisée, du « travail scientifique des données » (Borgman, 2020). En France, la recherche prend même part à la production des enquêtes publiques au côté des statisticien·ne·s et autres expert·e·s, au sein d'infrastructures publiques de production de connaissance associant des entités comme le CNRS (Centre national de la recherche scientifique), l'INED (Institut national d'études démographiques) ou l'INSEE (Institut national de la statistique et des études économiques). Par ailleurs, la communauté scientifique est elle-même de plus en plus poussée à penser les conditions de l'ouverture et de la « mobilité » des données qu'elle produit au travers des exigences de l'*open access*. De plus, leur inscription dans des structures publiques indépendantes est censée leur permettre d'échapper aux logiques d'influence, notamment des intérêts privés, et, par là, de garantir leur capacité critique. Pourtant, force est de constater que la recherche en sciences sociales est rarement associée aux démarches d'ouverture des données administratives.
- 9 Les données administratives sont d'abord produites pour outiller les politiques publiques et les actions des institutions. Ce sont des données de gestion, servant à organiser l'État — Alain Desrosières parle de « registres » (2005) — et qui sont aussi utilisées comme instruments de pilotage dans le contexte des différentes réformes néomanagériales dont les institutions publiques font l'objet (Bezès, 2005). C'est le cas des données comptables sur les subventions aux associations qui sont produites par la municipalité pour centraliser les informations et aider aux arbitrages financiers. Par exemple, à propos d'un fonds d'aide aux associations d'un million d'euros mis en place par la mairie au moment de la pandémie du COVID-19, l'élue déléguée à la vie associative explique en entretien :

On a fait une instruction collégiale [...] et on a regardé toute la trésorerie, toutes les finances des associations, et on a eu tous les cas de figure : des subventions données à des associations qui, si elles n'avaient pas ces 15 000 ou 20 000 euros, il y avait des risques qu'elles mettent la clef sous la porte [...] soit on a décidé de ne pas donner malgré des situations de COVID parce qu'on s'est rendu compte que l'association avait un an, deux ans de marge, bon on va se revoir plus tard ! [...] on a fait ce travail-là [...], vraiment dans la transparence, et ça nous a aidés à progresser je trouve et mieux connaître encore notre tissu associatif et nos partenaires.

- 10 On voit dans le discours de l'élue à quel point les données ont été centrales dans les arbitrages pour l'attribution des aides.
- 11 À une époque d'austérité généralisée, la question des financements publics dédiés aux associations est souvent taboue et sujette aux crispations. Dans un contexte d'évolution importante et de mise sous tension des modalités du soutien financier apporté au monde associatif par la puissance publique, on peut s'étonner que le premier, et notamment ses « têtes de réseaux », n'aient pas plus investi cet enjeu de l'*open data* — d'autant plus dans un contexte d'inflation des informations collectées auprès des associations par les administrations dans le cadre des procédures de financement. S'il est maintenant clair pour une grande partie des associations, a fortiori employeuses, que les financements publics tendent à baisser et que les conditions de leur obtention sont de plus en plus contraignantes, celles-ci ont pendant longtemps été dépourvues des ressources pour objectiver cette impression. C'est justement la recherche qui a été la première source de données vraiment solides, à savoir l'enquête nationale Paysage associatif (Tchernonog *et al.*, 2007, 2013 ; Tchernonog et Prouteau, 2019), aujourd'hui largement reprise par le secteur².
- 12 Tout en satisfaisant un intérêt propre à la connaissance scientifique du secteur, les chercheur·euse·s coordonné·e·s par Viviane Tchernonog ont fourni au monde associatif des indicateurs attestant de phénomènes tels que la baisse globale des financements publics, l'effacement progressif des subventions au profit de la commande publique ou encore l'augmentation de la part de financement par la participation des usager·ère·s et donc de la dépendance à la solvabilité de ces dernier·ère·s. Si cette enquête s'appuie sur des données publiques, produites notamment par l'INSEE, il est à noter qu'une grande part des informations, notamment concernant les réalités budgétaires, sont collectées par le biais d'un questionnaire administré par les chercheur·euse·s. En effet, les données administratives, notamment en ce qui concerne les financements publics, n'ont pendant longtemps pas existé, ou plutôt sont restées « fermées », difficiles d'accès ou pas directement exploitables.
- 13 Toutefois, la multiplication des initiatives publiques de diffusion de telles données montre que la situation évolue. On doit bien sûr mettre ce mouvement sur le compte des réglementations incitatives en faveur de l'*open data*, avec la création d'Etalab en 2011, puis la Loi pour une République numérique de 2016. En effet, en 2017, un décret³ puis un arrêté ministériel⁴ viennent définir les contours de l'obligation⁵ des collectivités locales de permettre l'accès aux données sur les « conventions de subvention » établies. Selon une disposition réglementaire datant de 2000⁶, la transparence sur les subventions n'est obligatoire que pour des conventions de subvention égales ou supérieures à 23 000 euros, un chiffre qui correspond aussi au seuil réglementaire de subventions à partir duquel la signature d'une convention est devenue obligatoire⁷. Or, au moment du lancement de notre recherche, de telles données n'étaient pas (encore) rendues accessibles. Faute d'être « données », ces données devaient donc être « obtenues » (Latour, 2007).
- 14 En effet, l'accès aux données a donc nécessité d'entrer en contact avec la municipalité. Leur obtention en quelques mois a été facilitée par une conjonction d'éléments favorables. Un premier élément tient à la stratégie partenariale que nous avons adoptée. Le projet de recherche n'était ni commandité ni financé par la collectivité étudiée. Tout l'enjeu était donc de montrer à ses représentant·e·s les bénéfices qu'ils et elles pouvaient tirer de cette main tendue de la recherche publique afin non seulement

d'obtenir l'accès aux données, mais aussi de construire un échange permettant de les comprendre. Afin d'obtenir un accord politique, nous avons présenté un projet ayant pour objectif « d'interroger les conditions financières de l'autonomie associative et plus particulièrement de l'accomplissement des missions démocratiques des associations (participation à la définition des politiques publiques, au débat public, interpellation, etc.) ». Nous précisons néanmoins : « Le contexte de diminution structurelle des dotations d'État et, par là, de la capacité budgétaire des collectivités tendent à avoir des effets importants sur les modalités de financement des associations qu'il convient d'objectiver par les méthodologies des sciences sociales. » L'idée était ici de ménager nos interlocuteur·rice·s, même de les mettre en confiance en leur montrant que notre démarche ne visait a priori pas à faire porter la responsabilité des difficultés des associations à la (seule) municipalité.

- 15 Outre la dimension politique, un deuxième élément de contexte favorable concerne une dimension plus technique. Lorsque notre demande intervient, à savoir près de quatre ans après les premiers décrets sur l'ouverture des données publiques, on peut supposer que les principes guidant cette ouverture ont été intégrés, voire banalisés, au sein de la municipalité, notamment à travers la mise en place de dispositifs et de procédures de rationalisation gestionnaire. Cette hypothèse se vérifie à travers nos premiers entretiens, notamment avec la personne responsable du contrôle de gestion des subventions. Pobin⁸ travaille à la ville depuis 2005 et a pendant 15 ans joué un rôle pivot dans les efforts de rationalisation des dispositifs de financement de la vie associative de cette ville. Nous ne reviendrons pas ici sur les différentes dynamiques que cette personne a été chargée de mettre en place, en s'appuyant au départ souvent sur des savoirs acquis pendant ses études de gestion, et donc notamment sur des normes provenant du monde de l'entreprise⁹. Cette facilité d'accès aux données peut notamment s'expliquer par l'outillage informatique. Comme le rappellent Denis et Goëta, « l'accessibilité des données est en effet toujours subordonnée aux logiciels qui les rendent visibles, les ordonnent » et « rares sont les bases de données qui sont affublées d'une fonctionnalité d'extraction automatique, qui permettrait de récolter les données indépendamment de leur mise en forme logicielle » (*op. cit.*, p. 25). En ce qui concerne l'ouverture des données municipales, certains instruments tels qu'un progiciel de gestion des subventions et des dispositifs informatisés de collecte d'informations ont été décisifs. Mis en place dans cette ville en 2010, Progos Aides versées est un progiciel édité par la société MGDIS. Au fil de la documentation fournie sur son site officiel, on trouve le leitmotiv « Reprenez le contrôle de vos financements » et des objectifs tournés vers la « maîtrise », l'« optimisation » ou le « pilotage intégré ». Pour autant, comme le montrent certaines recherches, les architectures informatiques mises en place dans une optique de standardisation gestionnaire et in fine d'efficacité font aussi souvent l'objet d'appropriations plus ou moins inattendues (Chiapello *et al.*, 2013 ; Nonjon et Marrel, 2015a, 2015b). Jamais, dans cette documentation grand public, la société ne met en avant les fonctionnalités propices à produire de la donnée réutilisable bien que le progiciel permette l'extraction des données dans un format ouvert. On voit ici comment la standardisation gestionnaire peut offrir des conditions propices à l'ouverture des données et servir in fine, là aussi, d'autres objectifs que l'efficacité organisationnelle et l'efficience budgétaire. La mise en place du progiciel a conduit non seulement à centraliser l'information provenant des différents services municipaux, mais aussi à la standardiser grâce à l'instauration de processus automatisés, de certains indicateurs et d'opérateurs communs.

- 16 Contrairement à une politique d'*open data* qui fonctionnerait par la prospection des données disponibles réalisée au sein même d'une administration, l'identification des données, l'une des « épreuves de diffusibilité » décrites par Courmont (2017), s'est donc ici faite à partir de notre demande, donc par une sollicitation extérieure¹⁰. Plus précisément, l'identification des données a eu lieu à la suite d'une chaîne d'échanges ayant débouché sur la rencontre avec un interlocuteur-clé, à savoir Pobin, contrôleur·euse de gestion ayant la main sur les données et la capacité de les extraire. Nous verrons par la suite que les échanges avec cette interlocuteur·rice ont aussi été déterminants dans la phase d'exploitation des données.

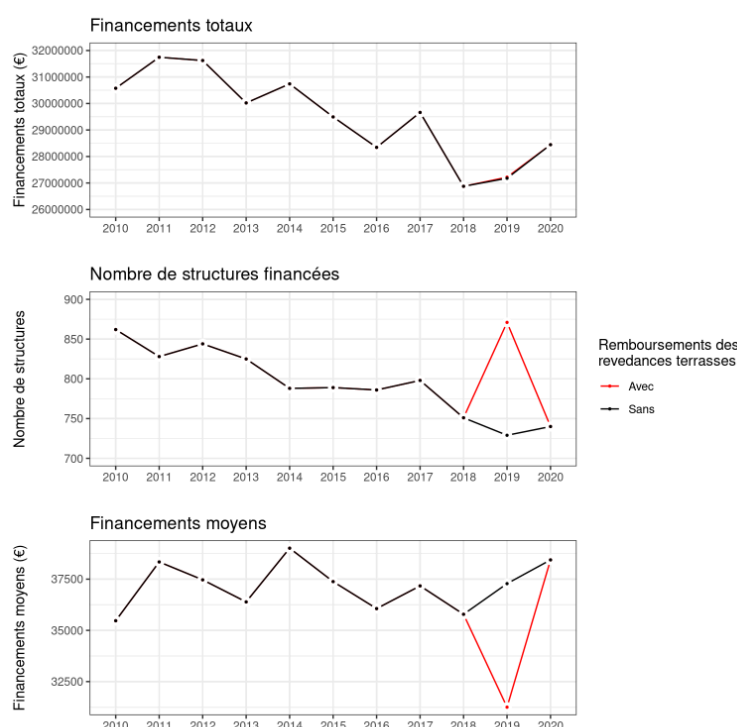
Des données modifiées : Contextualiser des données « métiers » pour les rendre exploitables pour la recherche

- 17 À ceux et celles qui conçoivent la donnée comme « un instrument de représentation objective d'une réalité qui lui préexiste [...] autonome par rapport à l'environnement social dans lequel elle est produite et utilisée », certain·es chercheur·euse·s comme Courmont rétorquent que « leurs usages initiaux sont inscrits dans le format et la structure des données, ce qui limite le champ d'utilisation possible de ces données » (2015, p. 43). Nous en avons fait l'expérience. L'extraction n'est en effet qu'une étape, certes décisive, mais non suffisante du processus d'ouverture des données. En ouvrant pour la première fois le fichier CSV extrait du progiciel par Pobin, nous nous sommes retrouvée·s face à des données comptables qui, sous plusieurs aspects, apparaissaient difficilement intelligibles. Pour qu'elles puissent être réutilisées, les travaux sur les politiques d'*open data* affirment que les données administratives, ou données « métiers », doivent faire l'objet d'un ensemble de transformations afin de les « détacher » de leur contexte initial (Labelle et Le Corf, *op. cit.*). Toutefois, le risque est alors de considérer les données comme dépouillées de leur contexte, de les mésinterpréter et de produire des contre-sens. Des recherches parlent de « consolidation » afin de permettre différents usages des données (Courmont, 2017) ou encore de « brutification » des données (Denis et Goëta, *op. cit.*), en s'appuyant sur la métaphore du pétrole et du processus de son raffinement, comme étapes d'épuration produisant des « “bonnes” données pour l'*open data* : des données au bon format, “propres” et intelligibles » (*ibid.*, p. 40).
- 18 Mais dans le cadre d'une réutilisation à des fins de recherche, à quel point la « désépécification » est-elle souhaitable ? Dans notre cas, fallait-il respecter l'intégrité des données telles qu'elles avaient été saisies ou bien les modifier et les actualiser en vue de mieux les comprendre ? Pour utiliser les données sur les subventions fournies par la municipalité dans l'optique de les mettre au service de nos questions de recherche, il nous est rapidement apparu qu'au lieu d'un travail de « désépécification », c'était bien à un travail de contextualisation qu'il était pertinent de s'atteler (Trupia, 2021). De notre point de vue, le travail de consolidation était, en fait, indissociable des intentions de l'usage. Ce travail de *consolidation par contextualisation* a alors consisté en des opérations à la fois de nettoyage et d'ajout d'indicateurs ouverts. Chaque opération de contextualisation s'est traduite par des phases de compréhension, puis de correction

des données. Chacune de ces étapes s'est faite en collaboration avec l'organe producteur des données.

- 19 Tel quel, le fichier fourni avait la forme d'un tableau avec des lignes correspondant à des opérations comptables positives et négatives. En plus des montants étaient renseignés le nom de la structure bénéficiaire, son numéro SIRET, l'année de financement, la direction municipale qui finance et une catégorie « objet ». Si cette dernière colonne est apparue peu standardisée et intelligible pour comprendre la nature des financements et leurs liens potentiels à des projets communs, certains termes comme *annulation* ou *remboursement* permettaient déjà de comprendre les montants négatifs. Toutefois, nous ne comprenions alors pas à quoi renvoyaient concrètement ces annulations : était-ce des récupérations de trop-perçus ? Des jeux comptables de répartition sur différentes années ? En faisant la somme des montants positifs et négatifs par structure, nous avons déjà pu obtenir les montants globaux obtenus par les structures financées, le nombre de ces dernières et l'évolution de ces deux dimensions sur la période, toutes directions et actions confondues. Par ces premières analyses exploratoires, plusieurs dimensions sont apparues problématiques et ont demandé des échanges avec l'administration productrice des données pour corriger ce qui nous apparaissait comme des aberrations.
- 20 D'abord, nous avons observé une hausse anormale du nombre de structures financées en 2019 associée à une hausse des subventions de petits montants. En regardant le détail, nous avons constaté qu'il s'agissait surtout de bars, cafés et brasseries. Dans un échange courriel, Pobin nous a finalement expliqué qu'il s'agissait d'une « exonération de redevances d'occupation du domaine public décidée par la collectivité suite aux manifestations des gilets jaunes » (graphique 1), en finissant par dire : « à voir comment vous les traitez, après tout il s'agit bien de subventions », ce qui est révélateur de l'ambiguïté des données : doit-on considérer cela comme une erreur à corriger ? La réponse peut être différente selon les usages et les finalités de l'exploitation des données.

Graphique 1. Montant des subventions avec et sans les remboursements des « redevances terrasses »



- 21 Les données fournies¹¹ concernaient l'ensemble des subventions, c'est-à-dire des sommes d'argent allouées par la municipalité à travers cette forme juridique et comptable. Ces sommes pouvaient donc être octroyées à d'autres entités que des associations.
- 22 En l'état, les données obtenues posaient une autre difficulté : les mêmes organisations pouvaient apparaître plusieurs fois lorsqu'elles avaient changé de numéro de SIRET pendant la période considérée, un changement qui survient notamment automatiquement lors d'un déménagement. Or, dans notre recherche, nous nous intéressions aux trajectoires de financement des structures, si bien que pouvoir retrouver ces dernières d'une année à l'autre à partir d'identifiants stables était un enjeu fort. La question de la mise à jour des données s'est révélée centrale. Les données « métiers » sont utilisées à l'instant T et il n'est pas nécessaire pour les organismes producteurs de garder un identifiant stable dans le temps. Pour notre étude, nous avons donc choisi de regrouper les doublons en les attachant au numéro SIRET le plus récent. Nous sommes donc intervenu·es sur les données et les avons modifiées afin de les rendre intelligibles dans le cadre de besoins de lecture pour le projet de recherche. Ce n'est qu'à ce moment-là que nous avons été capables, en nous appuyant sur des données consolidées, de décrire l'évolution du volume des structures financées et des montants annuels perçus par entité, ainsi que celle de l'enveloppe globale et des montants globaux octroyés par la municipalité sur la période. Il était alors possible de montrer que le nombre de structures financées par an diminuait au cours de la période, tout comme l'enveloppe globale des subventions qui leur étaient allouées, impliquant des financements moyens relativement stables (graphique 1). Ce travail sur les données montre qu'il serait par exemple pertinent pour les collectivités de construire un système de suivi historicisé des associations à travers l'agrégation de leurs numéros

SIRET, de même que de réaliser elles-mêmes les agrégats totaux de financements octroyés annuellement par structure.

- 23 Dans cette démarche de nettoyage, nous avons également cherché à rendre exploitables d'autres informations : par exemple, celle sur les directions thématiques ou services municipaux d'où proviennent les financements. Dans la table primaire, 60 intitulés de directions étaient renseignés. Outre le fait qu'un tel nombre de modalités ne soit pas propice à l'analyse, l'examen qualitatif des intitulés a montré qu'une harmonisation était nécessaire. D'abord, certaines formules étaient quasiment identiques, comme « direction des initiatives solidaires et de la jeunesse » et « direction des initiatives solidaires et jeunesse », ou encore « direction des solidarités et de la santé » et « direction santé-solidarité », induisant un besoin de standardisation. Ensuite, d'autres intitulés semblaient être des regroupements de deux autres : la « direction des initiatives solidaires et de la jeunesse » était sans doute un regroupement des « direction des initiatives solidaires » et « direction jeunesse » et suppose des réorganisations internes pendant la période. Nous avons d'ailleurs connaissance d'une réorganisation ayant débouché sur la création de la « direction de la proximité » à travers le regroupement de la « Politique de la ville » et de la « direction des quartiers ». De plus, la ville étudiée étant constituée de trois communes associées, les directions — lorsqu'elles apparaissaient dans les données — ne recoupaient pas les mêmes thématiques entre ces différentes communes. Comme notre exploration l'a montré, cet enjeu de stabilité temporelle et spatiale est un préalable incontournable à une étude de l'évolution des priorités thématiques de l'institution et de sa retranscription dans la structure des financements des associations. Nous avons donc entrepris de regrouper nous-mêmes les différentes directions, non seulement dans l'optique de créer une continuité spatiale et temporelle, mais aussi de créer des agrégats thématiques propices à l'étude (tableau 1). Là encore, le contact établi avec Pobin a été crucial et a permis de valider la pertinence de ces regroupements.

Tableau 1. Services municipaux regroupés en thématiques d'action

Thématique	Direction
Administration interne	Cabinet du maire
	Direction générale
	Services du maire
	[Commune associée 1] Direction de la population
	Direction de la production
	Relations sociales et du travail
	Direction générale ase
	Élections
	Direction foncière et immobilière

	Finances
	Contrôle de gestion
	Gestion administrative du personnel
	Élections démographie recensement
Culture	[Commune associée 1] Éducation culture
	Spectacle vivant et musiques
	Direction du patrimoine
	Direction littérature et lecture publique
	Direction des arts visuels
	Musée de géologie et houiller
	Tic
Animation territoriale	Politique de la ville
	Direction de la proximité
	Direction des quartiers
	Équipe direction politique de la ville
	Événementiel
	Logistique des manifestations
	Logistiques fêtes et manifestations
Cadre de vie	Gestion espace public cadre de vie
	Urbanisme aménagement
	Direction de l'espace public
	Qualité et développement de la ville
	Direction de projet renouvellement urbain
	[Commune associée 1] Pôle qualité développement de la ville
Jeunesse	Direction petite enfance
	Direction des initiatives solidaires et de la jeunesse
	Direction jeunesse – animation

	Direction des initiatives solidaires
	Direction jeunesse
	Direction des initiatives solidaires et jeunesse
Santé-solidarité	Direction personnes âgées handicapées
	Direction des solidarités et de la santé
	[Commune associée 1] Solidarité
	Direction santé-solidarité
	Ville de la solidarité
Éducation	Direction éducation – actions éducatives
	Direction éducation
	Direction éducation peg
	Direction éducation – service ressources
Développement économique	Économie
	Développement économique relations internationales
	Économie tourisme commerce
	Action économique commerce
Sécurité publique	Clspd
	Police municipale
Sports	Direction des sports
Associations	Maison des associations
Environnement	Développement durable
Habitat	Habitat risques urbains et sanitaires
Défense des droits	Maison de la médiation et du citoyen
[Commune associée 2]	[Commune associée 2]

24 Au travers de ces exemples, on comprend que la contextualisation des données est essentielle pour éviter des contre-sens ou erreurs, de même que pour envisager des usages pertinents en matière de recherche. Cette démarche illustre donc l'importance des échanges avec l'administration qui produit les données. Grâce à ce travail de consolidation des données sur les origines thématiques des financements, nous avons

fait apparaître les thématiques les plus financées comme la jeunesse, l'éducation ou la culture, ainsi que des thématiques plus anecdotiques comme l'environnement et la défense des droits. On entrevoit alors des pistes de recherche en termes de *path dependency* ou de changement des priorités politiques. Les associations sont souvent financées quasi exclusivement par un secteur thématique correspondant à un secteur d'action publique spécifique — comme la culture ou l'éducation.

- 25 La pertinence de ce travail d'articulation des directions est toutefois limitée par le fait que les données n'informent pas sur le cofinancement d'un même projet par plusieurs directions. Selon Pobin, dans ces cas, des « virements » ont même souvent lieu entre directions thématiques, notamment lorsqu'il s'agit d'un financement dépassant 15 000 euros, donc impliquant une convention : « c'est pas obligatoire, mais c'est très pratique ». Ainsi, le cofinancement et, avec lui, au moins l'une des directions thématiques impliquées disparaissent de la comptabilité, ce qui tend à biaiser le travail d'exploitation de ces données — là encore, il faut en être informé.
- 26 Enfin, nous avons aussi cherché à comprendre la logique des données renseignées dans la colonne « Objet » afin d'en induire des fils d'analyse. Rapidement, nous avons cependant réalisé que ces données n'étaient pas exploitables. Nous avons à un moment espéré pouvoir regrouper les lignes par projet ou déceler des indications à propos des actions financées, mais les contenus ne renseignaient souvent rien en ce sens et se cantonnaient à des numérotations variables. Dans le cadre d'une démarche d'ouverture de ces données et pour le travail que nous menons, cette information aurait été utile afin que nous puissions distinguer les lignes afférentes à des subventions de fonctionnement ou à certains projets plus spécifiques.
- 27 Dans notre travail, la consolidation a donc consisté, loin d'une « déspecification », en des opérations d'actualisation, de contextualisation et de documentation qui se sont révélées nécessaires pour rendre des données « métiers » intelligibles et exploitables par des acteur·rice·s extérieu·re·s à l'administration qui les avait produites — opérations consistant en la fusion et l'addition de lignes ou encore en la construction de regroupements thématiques stables dans le temps.

Dépasser les données ? Les voies empêchées de l'ouverture des données

- 28 Les données fournies par la mairie, même une fois corrigées et contextualisées, restaient limitées. En les exploitant, il est apparu qu'elles pouvaient être enrichies par leur articulation à d'autres sources complémentaires d'informations, et ce, à travers le numéro de SIRET, qui pouvait jouer le rôle de ce que Courmont nomme une donnée « pivot » (2015, p. 47). Nos besoins de recherche ouvraient alors autant de pistes qu'ils montraient les limites des politiques d'*open data* telles qu'elles sont actuellement conçues. La perspective de l'*open data* du point de vue de l'offre implique l'idée que les organismes producteurs des données devraient travailler à les « détacher » de leur contexte (Labelle et Le Corf, *op. cit.*). Derrière cette approche, il y a la fiction d'un état brut, autonome, voire purifié des données, dont pourrait s'approprier n'importe quel·le acteur·rice. Or, une approche relationnelle de l'*open data* nous pousse à reconnaître qu'il n'est pas vraiment possible, sinon pertinent de penser l'ouverture des données administratives, non seulement en dehors des contextes sociaux dans lesquels elles ont

été produites, mais aussi de ceux où elles sont appelées à être utilisées. Les données mises à disposition doivent être rendues intéressantes pour les utilisateur·rice·s. Même si le retour sur l'intérêt des données ouvertes n'en est pas l'objectif principal, certaines administrations ont mis en place des « hackathons », concours d'innovation à travers des projets de développement d'applications ou de services à l'appui de la réutilisation de données mises à disposition. Ces événements sont l'occasion d'avoir un retour sur le contenu de leurs données et permettent à ces administrations de réaliser que les données doivent être mises en forme ou complétées, que ce soit en amont (Denis et Marquet, 2019) ou durant les concours (Trupia, *op. cit.*). La mise à disposition des données est indissociable de la projection qu'elles permettent dans un horizon d'exploitation répondant aux préoccupations et intérêts des acteur·rice·s. Or, cette projection dépend aussi de la disposition des données à être enrichies, articulées à d'autres données, ce que nous montrons à travers certaines tentatives, plus ou moins fructueuses, réalisées dans le cadre de notre recherche. S'agissant de données comptables, les données municipales étudiées ne fournissaient quasiment aucune information sur les structures financées, qu'il s'agisse de leur forme juridique, de leur nombre d'employé·es ou de leur secteur d'activité. S'il pouvait sembler évident que l'ensemble des structures bénéficiaires de subventions associatives sont des associations, il s'est révélé que ce n'était pas le cas, comme nous avons pu le décrire à propos des remboursements des redevances terrasses. Nous avons donc dû trouver cette information au moyen d'autres sources.

- 29 Pour collecter plus d'informations, il a été possible, en nous appuyant sur le numéro SIRET — le plus récent pour chaque structure financée —, d'extraire les données contenues dans d'autres bases, notamment la base SIRENE. Celle-ci nous a alors permis de spécifier le statut juridique des organisations financées afin d'isoler les associations des autres types d'organisations. Là encore, les modalités disponibles ont été regroupées, ce qui nous a amené·es à distinguer les associations des entreprises, des organismes de l'économie sociale et solidaire et des institutions et administrations. Cela nous a permis d'objectiver le fait que la grande majorité des structures financées (91 %) par la municipalité par le biais de subventions sont bien des associations, cette part restant stable au cours de la période considérée.
- 30 L'articulation à la base SIRENE a également permis de savoir si les structures financées étaient employeuses ainsi que les tranches de nombre de salarié·es, permettant de distinguer les grandes associations délégataires de politiques publiques des petites associations susceptibles, plus ou moins régulièrement, de jouer le rôle de prestataires ou de porteurs de projets. Pour les plus grandes entités, les données SIRENE permettent même de savoir si elles sont composées de plusieurs établissements et si leur implantation dépasse le contexte communal (agglomération, région, pays), par exemple le CNRS et ses 908 établissements, la Croix-Rouge et ses 680 établissements ou encore APF France handicap qui en compte 559. En revanche, ce travail d'enrichissement est limité par les données disponibles au moment de la récupération des données sur les associations.
- 31 Ces données complémentaires nous ont aussi permis, à travers la NAF, nomenclature d'activités française, d'identifier les thématiques d'intervention des structures financées par la mairie. Les structures doivent se positionner elles-mêmes dans cette nomenclature. Le problème est qu'environ la moitié (46,5 %) des structures financées ont coché la modalité « Autres organisations fonctionnant par adhésion volontaire ».

Cela témoigne du fait que beaucoup d'associations ne se retrouvent pas dans cette nomenclature, qui serait mal adaptée pour typologiser les activités associatives.

- 32 Il existe bien une nomenclature des objets sociaux des associations adaptée à la nature de leurs activités dans le répertoire national des associations (RNA)¹², construite en adéquation avec les types d'activités des associations. Pourtant, dans les données fournies par la mairie, les structures financées étaient identifiées sur la base de leur numéro SIRET et non de leur identifiant RNA — toutes n'étant pas des associations. Mais les numéros RNA des associations ont été retrouvés à partir des numéros SIRET. Ainsi, alors que les 65 structures publiques ont été regroupées sous le libellé « Administration publique », nous avons, sur la base des descriptions de l'activité, reconstitué manuellement les secteurs d'intervention de 101 associations et 111 structures privées non appareillables au RNA¹³.
- 33 Peuvent ainsi être appréciées les orientations des financements qui se dirigent en priorité vers les secteurs culturel, de l'intervention sociale, de l'aide à l'emploi et du développement local ainsi que des sports (tableau 2). Il est cependant à noter que, si les structures culturelles sont les plus représentées à la fois en nombre de structures et en volume de financement (représentant respectivement 25,8 % des structures et 26 % du montant global), l'association parapublique City2000¹⁴ capte en réalité à elle seule 9,7 % des financements totaux. Cet exemple révèle les grandes disparités entre une structure de délégation de service public financée à hauteur de 3,14 millions d'euros entre 2010 et 2020 et 490 autres structures ayant reçu en moyenne 108 007 euros sur cette même période. Il en va de même pour deux autres structures qui ont reçu plus de 15 millions d'euros (environ 5 % du total) et ont été isolées dans le tableau 2 : l'organisme départemental de gestion des établissements de l'enseignement catholique représente 5,8 % des financements et l'organisme de gestion des missions locales, 5,1 %.

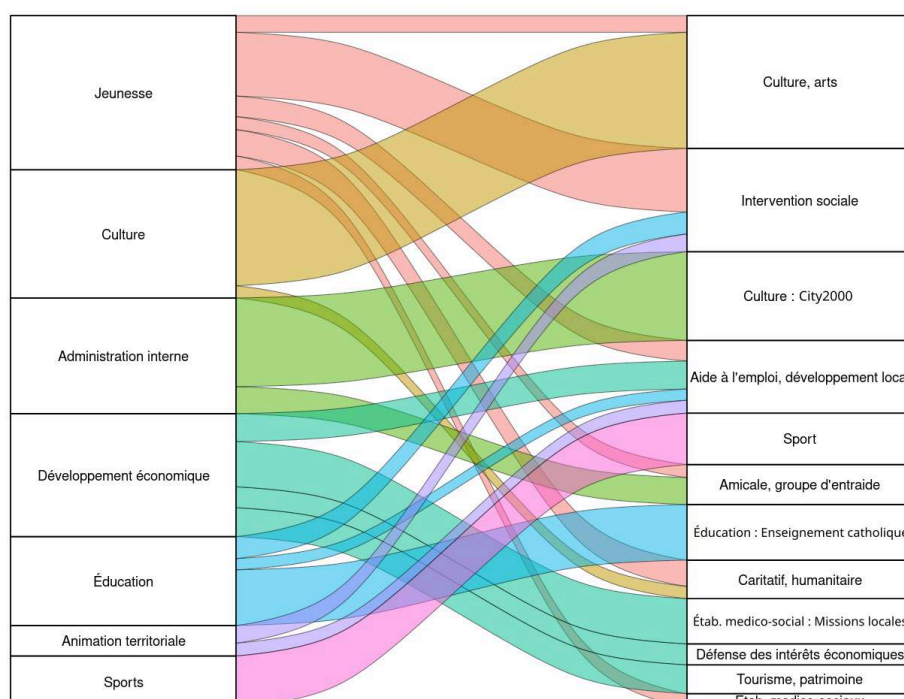
Tableau 2. Thématiques d'action des associations subventionnées selon la nomenclature RNA

Secteur d'intervention	Structures		Financement total (2010-2020)		Financement moyen (€)
	Nombre	Pourcentage (%)	Montant (€)	Pourcentage (%)	
Culture, arts (hors City2000)	490	25,8	52 923 282	16,3	108 007
Culture, arts : City2000	1	0,1	31 400 000	9,7	31 400 000
Sport	219	11,5	25 088 782	7,7	114 561
Éducation, formation, recherche (hors enseignement catholique)	140	7,3	6 216 487	1,9	44 403
Éducation, formation, recherche : enseignement catholique	1	0,1	18 755 445	5,8	18 755 445

Intervention sociale	108	5,7	39 750 334	12,2	368 059
Amicale, groupe d'entraide	93	4,9	18 855 258	5,8	202 745
Défense des intérêts économiques	94	4,9	12 655 347	3,9	134 631
Caritatif, humanitaire	81	4,3	18 698 488	5,8	230 846
Santé	80	4,2	2 341 230	0,7	29 265
Action socio-culturelle	78	4,1	4 585 775	1,4	58 792
Club de loisir	68	3,6	5 159 399	1,6	75 874
Administration publique	67	3,5	2 483 407	0,8	37 066
Aide à l'emploi, développement local	61	3,2	29 024 800	8,9	475 816
Environnement, cadre de vie	58	3,1	9 307 540	2,9	160 475
Défense des droits	41	2,2	3 130 438	1	76 352
Établissements médico-sociaux (hors missions locales)	39	2	16 596 708	5,1	425 557
<i>Établissements médico-sociaux : missions locales</i>	1	0,1	9 965 446	3,1	9 965 446
Activité économique	35	1,8	1 431 528	0,4	40 901
Information, communication	32	1,7	728 682	0,2	22 771
Tourisme, patrimoine	26	1,4	10 401 530	3,2	400 059
Logement	22	1,2	3 519 579	1,1	159 981
Sécurité, armée, justice	11	0,6	76 120	0	6 920
Religion	6	0,3	350 325	0,1	58 388
Politique, cercle de réflexion	5	0,3	8 500	0	1 700
NR	40	2,1	1 249 987	0,4	31 250
Ensemble	1897	100	324 704 414	100	171 167

- 34 En plus de ces orientations globales, l'articulation des orientations thématiques des directions municipales et des structures financées offre une perspective plus ciblée des fléchages donnés aux financements au sein de chaque direction. Ainsi, alors que l'association City2000 intervient dans le secteur de la culture, son financement provient de la direction thématique « Administration interne » et occupe par ailleurs une part importante de l'enveloppe de cette direction. Pour les autres, se détachent des secteurs d'intervention largement financés par une seule thématique de direction comme la culture (hors City2000) et les sports, et d'autres trouvant des financements plus éclectiques à la mairie comme l'aide à l'emploi ou l'intervention sociale (graphique 2).

Graphique 2. Origines et destinations thématiques des subventions associatives de la mairie entre 2010 et 2020



Champ : Transferts des subventions aux associations entre les thématiques de direction et les secteurs d'intervention correspondant à au moins 1 % du montant global.

- 35 Par ailleurs, les recherches existantes montrent qu'il est rare qu'une association employeuse vivant de subventions ne dépende que d'un seul financeur (Tchernonog et Prouteau, *op. cit.*). Le plus souvent, les associations sont financées non seulement par plusieurs institutions (collectivités territoriales, État, CAF, Pole emploi, etc.), mais aussi parfois par plusieurs services au sein de ces institutions (Chevallier, 2022). Ces financements multiples sont souvent dédiés à des projets communs, si bien qu'on les nomme des « cofinancements ». On a déjà évoqué plus haut que les cofinancements entre plusieurs services au sein de la mairie ne sont pas apparents dans les données : la simplification comptable conduit à l'affichage d'une seule direction et à la disparition d'autres, constituant un important angle mort des données municipales.
- 36 Un autre enjeu de notre recherche a donc été de dépasser les seules données sur les subventions municipales, conduisant par là même à une interrogation sur

l'interopérabilité des données administratives. Sur notre sollicitation, Pobin nous a procuré un complément de ce qu'il appelait les « cofinancements », disponibles uniquement pour les structures financées par la ville à plus de 23 000 euros et sur une période moins longue. Ainsi, ce filtrage mettait déjà de côté un nombre important de structures auxquelles les montants alloués restaient, pour une part, substantiels et pouvaient quand même être utilisés pour payer des salaires. Mais surtout, ce qui était nommé les « cofinancements » correspondait en fait au budget global des structures, ne permettant de produire qu'un pourcentage global de dépendance à la mairie et pour seulement environ un dixième des structures.

- 37 Pour compléter ces informations, nous nous sommes donc tourné·es vers les plateformes et offres d'*open data* des autres financeurs présents sur le territoire aux différents échelons. En France, les dispositions réglementaires évoquées plus haut ont certes incité les différentes collectivités à mettre en place les procédures nécessaires à l'ouverture des données sur les subventions, en même temps que beaucoup d'autres, par le biais de portails dédiés. Mais ces données sont pour l'instant restées largement sous-exploitées. Sur le territoire de notre étude, la municipalité s'est associée au projet de portail de l'intercommunalité. Pendant l'enquête, les données sur les subventions de cette municipalité ont d'ailleurs été publiées, mais seulement concernant les conventions de subvention de plus de 23 000 euros, correspondant au seuil réglementaire obligatoire.
- 38 En ce qui concerne les données des autres échelons administratifs du territoire, nous avons rapidement fait face à plusieurs obstacles. D'abord, tous les financeurs pertinents ne mettent pas à disposition de telles données, à l'instar du Conseil départemental ou de la Région. Ensuite, même dans le cas des données rendues disponibles par l'intercommunalité¹⁵ ou par l'État¹⁶, la variabilité des contenus des données et périodes d'observation limite la possibilité de les articuler aux données municipales, illustrant la grande hétérogénéité de production et de gestion de données de même nature. Cette articulation pourrait être rendue possible par un travail de dialogue et de coopération, notamment en vue de la mise à disposition de données qui ne sont pas encore rendues disponibles, et par un travail de nettoyage et d'harmonisation des données rendues disponibles par les différentes institutions. Néanmoins, on peut anticiper qu'un obstacle à ce travail serait l'existence de dispositifs administratifs et comptables très différents et inégaux quant à leur capacité à produire de la donnée, a fortiori sur une longue période.
- 39 Toujours est-il que toutes ces difficultés soulignent la nécessité d'un dialogue entre des institutions à différents échelons, dans l'optique d'avoir la capacité de produire des données en partie standardisées, qui puissent être articulées entre elles. Au sein de l'État, l'initiative Data.subvention aurait pu aller dans ce sens. Cette initiative est portée par la plateforme Beta.gouv.fr, un « incubateur de services numériques » visant à promouvoir et à accompagner, au sein des administrations, le développement de « start-ups d'État » et d'« APIs » (*Application Programming Interface*). Sur le site de présentation du projet Data.subvention, il est notamment indiqué que,
- face à l'importance du nombre de subventions versées chaque année et face à la volonté d'accroître le niveau de transparence concernant l'usage de l'argent public, il est essentiel aujourd'hui d'appliquer une politique d'ouverture des données, rendant disponibles et publiques, l'ensemble des subventions versées par l'État aux associations¹⁷.

- 40 Mais d'autres enjeux sont également évoqués autour de l'amélioration de l'information dans le cadre de l'octroi de subventions. Bien que des enjeux soient soulevés pour les associations (notamment le « non-recours par manque d'information ou par renoncement »), il semble que les solutions proposées soient surtout conçues en direction des services de l'État :

chaque service public financeur déploie du temps et de l'énergie à communiquer et informer sur ses dispositifs de subvention, à accompagner les associations demandeuses, à opérer des vérifications administratives et réglementaires et à relancer les porteurs de projets quand les dossiers sont incomplets ou non conformes, provoquant de nombreux allers-retours. Pour autant, aucun service n'a aucune vision globale et partagée de la situation d'une association.

- 41 Pour l'heure, le projet consiste à collecter les données sur les subventions auprès de différents services de l'État, mais aussi auprès de certaines collectivités, et de les mettre en commun dans le cadre d'une plateforme exclusivement accessible aux services de l'État. À ses débuts en 2022, notre équipe de recherche, pourtant inscrite dans un laboratoire public de recherche, s'est vu refuser son accès. On est donc, pour l'instant, loin d'une démarche visant la production de données ouvertes. En effet, cette initiative a été confiée à une « start-up d'État » et, à aucun moment, des formes de collaboration avec la recherche n'ont été envisagées. D'après nos échanges avec différentes personnes dans les principales têtes de réseaux du monde associatif, ces dernières sont elles-mêmes gardées à distance de la démarche. Malgré les discours (obligés) sur l'ouverture, cette initiative est donc visiblement surtout motivée par des objectifs gestionnaires de rationalisation de l'octroi des subventions dans un contexte d'austérité.

Conclusion : l'ouverture des données administratives, avec ou sans la recherche ?

- 42 Le travail scientifique d'exploitation de données administratives sur les subventions dédiées aux associations permet de repenser les démarches d'*open data*. Si ces dernières tendent à être largement pensées selon une logique d'offre, et donc à rester le plus souvent des « données » qui ne sont jamais « saisies » (Becker, *op. cit.*), notre expérience montre qu'elles ont tout intérêt à intégrer la demande. Pour que les données soient effectivement exploitées, il convient pour leurs organismes producteurs d'ouvrir un dialogue avec les actrices et acteurs sociaux intéressés afin d'adapter au mieux leur offre, selon leurs possibilités et leurs contraintes. Ce renversement de l'approche vers la demande induit aussi de ne plus considérer que de « bonnes » données soient des données « brutifiées » ou « déspécifiées » (Courmont, 2017 ; Denis et Goëta, *op. cit.*), mais plutôt des données contextualisées, c'est-à-dire que leurs organismes producteurs doivent en fournir les clés de compréhension, voire qu'elles doivent être transformées pour être rendues lisibles et intelligibles par des acteur·rice·s non spécialistes.
- 43 Dans les démarches d'*open data*, la recherche publique peut, par ses ressources propres (les compétences, le temps disponible, la position tierce), jouer un rôle d'intermédiaire, sinon de traduction. Si elle part forcément d'intérêts propres, elle construit ses réflexions à partir d'une connaissance des questionnements des acteur·rice·s concerné·es. Ainsi, l'intervention de la recherche dans l'ouverture des données peut être précieuse afin de rendre les données intéressantes pour les acteur·rice·s

impliquée-s tout en les sensibilisant à l'enjeu même de leur ouverture. En effet, un tel enjeu est apparu au cours de plusieurs moments de restitution de notre recherche impliquant des membres des sphères universitaire, institutionnelle et associative. Du côté du monde associatif, les acteur-ric-e-s ont exprimé un véritable enthousiasme face à la démarche, aux résultats et aux possibilités ouvertes en termes de plaidoyer. Dans le cas de la municipalité étudiée, la restitution des résultats a également permis non seulement de confirmer certaines observations faites en interne, mais aussi d'ouvrir certains débats sur le contexte actuel de baisse des subventions et le rôle potentiel des données administratives dans l'amélioration des rapports entre pouvoirs publics et associations. Notre cas illustre donc comment la recherche peut accompagner à la fois les administrations dans leur démarche d'ouverture de leurs données et la sphère associative dans ses connaissances et rapports avec les institutions. En partant de la connaissance des préoccupations des acteur-ric-e-s qu'ils et elles enquêtent, les chercheur-euse-s peuvent relayer ces dernières auprès des pouvoirs publics, ouvrir des voies d'exploitation des données et ainsi susciter une demande sociale et militante, et y répondre (Bruno, Didier et Prévieux, *op. cit.*).

- 44 Dans notre cas, il n'est pas sûr que les acteur-ric-e-s du monde associatif aient le temps, les compétences ou encore l'envie de produire des analyses à partir des données sur les subventions. Ces associations sont pourtant très demandeuses de telles analyses, et la recherche peut, comme c'est le cas pour l'enquête Paysage associatif (Tchernonog et Prouteau, *op. cit.*), assumer ce rôle. Mais, pour que la recherche puisse jouer ce rôle d'intermédiaire, il faut que lui soit reconnue par les institutions une place dans les politiques d'*open data*. Or, on peut regretter que les institutions semblent avoir moins tendance à se tourner vers la recherche publique que vers des acteur-ric-e-s inscrites dans les rationalités managériales, voire marchandes, et faisant miroiter des promesses de valorisation économique.

BIBLIOGRAPHIE

BAUDOT Pierre-Yves (2015), « Introduction », *Informations sociales*, 191(5), p. 4-7.

BAUDOT Pierre-Yves, Guillaume MARREL et Magali NONJON (2015), « Encore une révolution informatique ? Open et big data dans les organisations administratives », *Informations sociales*, 191(5), p. 8-18.

BECKER Howard (1952), « Science, culture, and society », *Philosophy of Science*, 19(4), p. 273-287.

BEZES Philippe (2005), « Le modèle de "l'État-stratège" : genèse d'une forme organisationnelle dans l'administration française », *Sociologie du travail*, 47(4), p. 431-450.

BORGMAN Christine L. (2020), *Qu'est-ce que le travail scientifique des données ? Big data, little data, no data*, Marseille, OpenEdition Press.

BOWKER Geoffrey C., Karen BAKER, Florence MILLERAND et David RIBES (2009), « Toward information infrastructure studies: Ways of knowing in a networked environment », dans Jeremy

- HUNSINGER, Lisbeth KLAstrup et Matthew ALLEN (dir.), *International Handbook of Internet Research*, Dordrecht, Springer Netherlands.
- BRUNO Isabelle, Emmanuel DIDIER et Julien PRÉVIEUX (dir.) (2014), *Statactivisme : comment lutter avec des nombres*, Paris, Zones Éditions.
- CHEVALLIER Thomas (2022), « Financements publics et limitation de l'autonomie des associations dans les quartiers populaires. Une démarche exploratoire par ethnographie budgétaire auprès de deux associations à Lille », *Sociologie*, 13(4), p. 439-459.
- CHIAPELLO Ève, Patrick GILBERT, Céline BAUD, Marion BRIVOT, Carine CHEMIN-BOUZIR, Corine EYRAUD et Bénédicte GRALL (2013), *Sociologie des outils de gestion : introduction à l'analyse sociale de l'instrumentation de gestion*, Paris, La Découverte.
- CONRADIE Peter et Sunil CHOENNI (2014), « On the barriers for local government releasing open data », *Government Information Quarterly*, 31, p. S10-S17.
- COURMONT Antoine (2015), « Open data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique », *Informations sociales*, 191(5), p. 40-50.
- COURMONT Antoine (2017), « De la “donnée” à la “donnée ouverte” : les épreuves de l'ouverture des données », *Statistique et société*, 5(3), p. 19-23.
- DENIS Jérôme et Samuel GOËTA (2017), « La fabrique des données brutes : le travail en coulisses de l'open data », dans Clément MABI, Jean-Christophe PLANTIN et Laurence MONNOYER-SMITH (dir.), *Ouvrir, partager, réutiliser : Regards critiques sur les données numériques*, Paris, Éditions de la Maison des sciences de l'homme, DOI : <https://doi.org/10.4000/books.editionsmsmh.9050>.
- DENIS Jérôme et Clément MARQUET (2019), « De la production à la maintenance : faire exister collectivement des données géographiques ouvertes », *Questions de communication*, 36(2), p. 63-83.
- DESROSIERES Alain (2005), « Décrire l'État ou explorer la société : les deux sources de la statistique publique », *Genèses*, 58(1), p. 4-27.
- DOMARD Julien (2012), « Le rôle de la mobilisation dans l'ouverture des données publiques : le cas de la publicisation de la localisation des parcelles transgéniques », *Les Enjeux de l'information et de la communication*, 13(2), p. 47-58.
- EDWARDS Paul N., Matthew S. MAYERNIK, Archer L. BATCHELLER, Geoffrey C. BOWKER et Christine L. BORGMAN (2011), « Science friction: Data, metadata, and collaboration », *Social Studies of Science*, 41(5), p. 667-690.
- GOËTA Samuel et Clément MABI (2014), « L'open data peut-il (encore) servir les citoyens ? », *Mouvements*, 79(3), p. 81-91.
- JANSSEN Marijn, Yannis CHARALABIDIS et Anneke ZUIDERWIJK (2012), « Benefits, adoption barriers and myths of open data and open government », *Information Systems Management*, 29(4), p. 258-268.
- LABELLE Sarah et Jean-Baptiste LE CORF (2012), « Modalités de diffusion et processus documentaires, conditions du “détachement” des informations publiques. Analyse des discours législatifs et des portails open data territoriaux », *Les Enjeux de l'information et de la communication*, 13(2), p. 209-259.
- LATOUR Bruno (2007), « Pensée retenue, pensée distribuée », dans Christian JACOB (dir.), *Lieux de savoir, 1. Espaces et communautés*, Paris, Albin Michel, p. 605-615.

- MARTIN Chris (2014), « Barriers to the open government data agenda: Taking a multi-level perspective », *Policy & Internet*, 6(3), p. 217-240.
- NONJON Magali et Guillaume MARREL (2015a), « Gouverner par les architectures informatiques. Logiciels et progiciels de gestion intégrée dans le secteur social », *Gouvernement et action publique*, 4(2), p. 9-24.
- NONJON Magali et Guillaume MARREL (2015b), « Le progiciel socialisé. Usages des architectures informatiques dans l'aide et l'action sociales départementales », *Gouvernement et action publique*, 4(2), p. 129-156.
- OLLION Étienne et Alexandre HOBEIKA (2018), « Vers une "data-sociologie" ? Tour d'horizon des techniques de collecte automatisée de données sur le web », blogue QUANTI/Sciences sociales, <https://quanti.hypotheses.org/647>, page consultée le 18 septembre 2025.
- PAQUIENSÉGUY Françoise et Valentyna DYMYTROVA (2018), « Open data et métropoles, les enjeux d'une transformation à l'œuvre. Analyse sémio-pragmatique d'un corpus de portails métropolitains », *Questions de communication*, 34(2), p. 209-228.
- RICŒUR Paul (1986), *Du texte à l'action : essais d'herméneutique*, Paris, Seuil.
- RUIJER Erna, Stephan GRIMMELIKHUIJSEN, Jochem VAN DEN BERG et Albert MEIJER (2020), « Travailler avec les données ouvertes : comprendre l'utilisation des données ouvertes dans une optique pratique », *Revue internationale des sciences administratives*, 86(1), p. 5-21.
- TALPIN Julien et Pierre BONNEVALLE (2023), « Financements associatifs et pouvoir local Enquête sur les subventions aux associations dans une ville du Nord », *Gouvernement et action publique*, 12(2), p. 37-64
- TCHERNONOG Viviane, Henry NOGUÈS et Muriel TABARIÉS (dir.) (2007), *Le paysage associatif français 2007 : mesures et évolutions : profil, activités, budget, financement, dirigeants, gouvernance, emploi salarié, travail bénévole*, Paris, Juris Éditions.
- TCHERNONOG Viviane et Lionel PROUTEAU (dir.) (2019), *Le paysage associatif français : mesures et évolutions*, Paris, Dalloz.
- TCHERNONOG Viviane, Lionel PROUTEAU, Muriel TABARIÉS, Henri NOGUÈS, Valérie FOURNEYRON et Hugues SIBILLE (2013), *Le paysage associatif français : mesures et évolutions*, Paris, Juris Éditions.
- TRUPIA Dilara Vanessa (2021), « Open transport data et développement d'applications de mobilité : Un travail d'équipement à la frontière de mondes variés », *Réseaux*, 4(228), p. 95-129.

NOTES

1. Nous remercions Julien Talpin, qui a participé à la mise en place de l'enquête sur laquelle cet article porte et avec qui nous avons échangé sur son contenu, ainsi que les élues et technicien·nes de la collectivité étudiée. Nous remercions aussi Clément Mabi, qui a pris le temps de nous faire des retours sur ce travail en relisant l'article.
2. Parmi de nombreux exemples, voir *L'état des assos en 12 infos. Repères sur les associations et le bénévolat en France*, rapport publié par Le Mouvement associatif en 2020.
3. Décret n° 2017-779 du 5 mai 2017 relatif à l'accès sous forme électronique aux données des conventions de subvention.
4. Arrêté du 17 novembre 2017 relatif aux conditions de mises à disposition des données essentielles des conventions de subvention.

5. Visant l'application de l'article 18 de la Loi pour une République numérique (n° 2016-1321 du 7 octobre 2016).
 6. Décret no 2001-495 du 6 juin 2001 pris pour l'application de l'article 10 de la loi n° 2000-321 du 12 avril 2000 et relatif à la transparence financière des aides octroyées par les personnes publiques.
 7. Circulaire dite « Valls », n° 5811-SG du 29 septembre 2015 relative aux nouvelles relations entre les pouvoirs publics et les associations.
 8. Nous faisons le choix de présenter Pobin de façon dégenrée dans ce texte. Le genre de cette personne n'a pas d'incidence pour les analyses que nous fournissons dans l'article.
 9. Nous travaillons sur un autre article sur le gouvernement par la rationalisation gestionnaire des modalités de subvention.
 10. Le fait que les données concernant les conventions de subvention de plus de 23 000 euros aient été rendues accessibles sur le portail *open data* pendant l'enquête — nous y revenons plus bas — donne d'ailleurs à penser que notre démarche a contribué à les encourager en ce sens.
 11. L'ensemble des résultats présentés dans l'article — tableaux et graphiques — a pour source les données sur les subventions de la mairie étudiée (2010-2020) et est le produit de traitements des auteur·rices.
 12. Fichier national répertoriant les informations sur l'ensemble des associations en France, produit par le ministère de l'Intérieur.
 13. Les « objets sociaux parents », regroupement plus large de secteurs d'activité, ont été pris en compte. Nous avons regroupé plusieurs secteurs comme suit : « Activités politiques » et « Clubs, cercles de réflexion » en « Politique, cercle de réflexion » ; « Sécurité, protection civile », « Armée (dont préparation militaire, médailles), anciens combattants » et « Justice » en « Sécurité, armée, justice » ; « Éducation, formation » et « Recherche » en « Éducation, formation et recherche » ; et « Tourisme » et « Préservation du patrimoine » en « Tourisme et patrimoine ». Les libellés des autres secteurs ont pu être raccourcis.
 14. La structure que nous avons renommée City2000 gère les financements de la ville pour l'organisation de la programmation de grands événements artistiques et culturels bisannuels à l'échelle de l'intercommunalité dans la poursuite d'un grand événement culturel européen.
 15. Disponibles pour la période 2015-2020, les données sur les subventions versées aux associations par l'intercommunalité ne sont pas limitées à des montants perçus minimaux, notamment au seuil de 23 000 euros, comme on peut le trouver dans les données similaires des communes de la métropole, mais n'intègrent ni objets de financement ni origines en termes de direction ou de thématiques.
 16. Les subventions venant de l'État apparaissent dans les annexes « jaunes » des projets de loi de finance. Il est possible d'harmoniser les contenus intégrant des objets de financement à partir de 2016.
 17. Voir <https://beta.gouv.fr/startups/data-subvention.html>.
-

RÉSUMÉS

Cet article montre comment le travail scientifique d'exploitation de données administratives sur les subventions aux associations permet de repenser les démarches d'*open data*. Si celles-ci sont essentiellement conçues selon une logique d'offre, notre expérience conduit à défendre une

démarche d'ouverture des données tournée vers la demande, qui intègre en amont ses utilisateur·rice·s potentiel·le·s. Ce renversement de l'approche vers la demande induit aussi de ne plus considérer que de « bonnes » données doivent être « déspecifiées », mais plutôt contextualisées. Enfin, nous montrons que la recherche publique peut jouer un rôle précieux d'intermédiaire en accompagnant les administrations dans leur démarche d'ouverture de leurs données ainsi que les actrices et acteurs sociaux dans leur appropriation.

This article demonstrates how the scientific work of exploiting administrative data on grants to associations can help rethink open data approaches. While these are primarily designed according to a supply-driven logic, our experience advocates for an open data approach focused on demand, one that integrates potential users at the outset. This shift toward a demand-driven approach also leads to the conclusion that “good” data should not be generalized, but rather contextualized. Lastly, we show that public research can play a valuable role as an intermediary by supporting administrations in their efforts to open up their data and social actors in their appropriation of it.

Este artículo se refiere al hecho de cómo el trabajo científico de utilización de los datos administrativos en las subvenciones a las asociaciones permite replantear los procedimientos de datos abiertos (*open data*). Si estos han sido esencialmente concebidos según una lógica de la oferta, nuestra experiencia da lugar a defender un método de apertura de los datos orientados hacia la solicitud, que integra con anterioridad a sus usuarios potenciales. Este giro en el enfoque hacia la solicitud también impulsa a no considerar solamente que los “buenos” datos deben ser “despecificados”, sino más bien contextualizados. Finalmente, demostramos que la investigación pública puede jugar un rol valioso de intermediario, acompañando tanto a las administraciones en su proceso de apertura de datos, así como, a los actores sociales en su apropiación.

INDEX

Keywords : associations, administrative data, open data, funding, supply and demand, grants

Mots-clés : open data, associations, subventions, financements, données administratives

Palabras claves : asociaciones, datos administrativos, datos abiertos, financiamiento, oferta y solicitud, subvenciones

AUTEURS

CÉCILE RODRIGUES

Cécile Rodrigues, ingénieure d'études CNRS, CERAPS. Courriel : cecile.rodrigues@cnrs.fr

THOMAS CHEVALLIER

Thomas Chevallier, Chargé de recherche postdoc au FNRS rattaché à l'UCLouvain (laboratoire CriDIS) / Chercheur associé au CERAPS, thomas.chevallier@uclouvain.be

Cultures de la quantification et quantification de la culture

Les effets de la mise en nombre des pratiques littéraires sur les utilisateurs de plateformes numériques

Juliette Parmentier

- 1 La recherche que nous proposons ici de développer part d'un constat simple : la culture — et plus précisément la littérature — n'échappe pas aux logiques de « mise en nombre » qui traversent l'ensemble des sociétés contemporaines. En d'autres termes, « ce qui, auparavant, était exprimé seulement par des mots et non par des nombres » existe désormais « sous une forme numérique » (Desrosières, 2008, p. 10). Sur Internet, les pratiques littéraires sont donc représentées et promues au travers d'une multitude d'artefacts quantifiés aussi hétéroclites que des compteurs, des statistiques, des classements, des notes, des pourcentages, des graphiques, etc. — autant d'éléments qui participent de ce que nous appelons les « cultures de la quantification » et dont nous voulons ici analyser les effets.
- 2 Comme nous tenterons de le démontrer, ces artefacts quantifiés — qui sont sous-tendus par une conception très spécifique de ce que lire veut dire — divisent implicitement les pratiques littéraires entre celles qui sont *prises en compte* dans les statistiques et les autres qui, comme la lecture de bandes dessinées ou la relecture, *comptent* donc *moins* ou *pas*. Or, comme souvent, la genèse de ces nombres — les calculs et les choix ayant présidé à leur conception, et in fine la conscience de leur contingence — s'efface avec leur diffusion dans l'espace numérique : ils sont présentés puis perçus comme neutres et objectifs. Pour beaucoup de lecteurs, le nombre indiqué sur leur compteur Babelio ou Goodreads *devient* alors leur nombre de lectures, partagé comme tel avec leurs proches ou dans leurs vidéos YouTube, sur leurs blogues, dans leurs publications Instagram. Ce « nombre de lectures » — construction statistique couramment mobilisée par les lecteurs en contexte numérique — est mis en avant, esthétisé, discuté, négocié et optimisé, mais aussi critiqué comme symptôme de l'immixtion de dynamiques productivistes au sein d'un champ littéraire toujours dépeint comme un « sanctuaire à l'abri de la surveillance cynique des réseaux sociaux¹ » et de leurs logiques commerciales (Murray, 2021).

- 3 De fait, les chiffres portent en eux un idéal de croissance. Ils tendent vers un « *plus* » indéfini (Grosser, 2014) qui trouve dans les discours prosélytes des lecteurs un terreau fertile : parce que les bienfaits de la lecture ne sont jamais remis en question et que sa démocratisation fait l'objet de tant de soins — en somme, parce qu'il s'agit de « la seule activité dont on postule l'universalité au point de s'inquiéter du nombre des non-pratiquants » (Barth-Rabot, 2023, p. 10) —, l'inflation des pratiques littéraires est présentée comme bonne en soi. Notre objectif est ainsi de souligner la façon dont lire *plus* est devenu un objectif à part entière pour beaucoup de lecteurs, qui doivent toutefois veiller à ne pas lire *trop*. La mise en nombre des pratiques littéraires a en effet ceci de particulier que les lecteurs ne peuvent jamais célébrer trop ouvertement la croissance de leurs statistiques — ce qui serait perçu comme un rapport déplacé à la littérature, compétitif, productiviste ou consumériste. En somme, nous souhaitons montrer comment la quantification des pratiques littéraires favorise des discours et des pratiques souvent contradictoires, incitant les lecteurs à faire croître leurs métriques tout en manifestant une indifférence à leur égard, invitant à se prendre pour objet d'étude tout en déjouant les présomptions de narcissisme, à augmenter la cadence tout en réaffirmant la prévalence du plaisir.
- 4 Afin de mieux saisir le rapport que les lecteurs entretiennent à ces artefacts quantifiés, nous avons conduit 31 entretiens semi-dirigés² avec des lecteurs francophones inscrits sur deux plateformes littéraires (Babelio et Livraddict) et/ou publiant du contenu littéraire sur Instagram.
- 5 Comme nous le montrerons, le rapport que les individus entretiennent à la quantification de leurs pratiques littéraires est émaillé de contradictions qui reflètent bien les injonctions paradoxales auxquelles ils sont soumis et qui ne sont pas étrangères à la « place très particulière du livre » au sein du champ culturel (Gilliotte, 2022). Lise³, par exemple, a répété plusieurs fois au cours de notre entretien qu'elle « n'aime pas les chiffres », ajoutant que « pour [elle], les chiffres et la lecture n'ont rien à voir », mais indique dans la biographie de son compte Instagram le décompte de ses lectures de l'année (« 2025 : 16 livres, 21 BD et 2 mangas ») ; Alain⁴, qui participe à des *challenges* de lecture sur Babelio, défend l'idée qu'il s'agit uniquement d'un « défi contre [lui]-même », avant de s'insurger contre les participants qui gonflent leurs statistiques en trichant pour gagner. Ces dissonances, dont nous pourrions multiplier les exemples, illustrent bien la pervasivité de la quantification : parce qu'ils y sont sans cesse exposés, les utilisateurs les plus réticents sont souvent contraints d'adopter eux aussi la grammaire chiffrée des contenus littéraires sur Internet.
- 6 Dès lors, et après avoir mis en évidence que la question spécifique de la mise en nombre des pratiques littéraires trouve étonnamment peu d'écho dans la littérature scientifique, nous tenterons de caractériser les spécificités de la quantification lorsqu'elle s'applique à la pratique culturelle « la plus hautement symbolique et la plus investie de valeurs » (Barth-Rabot, *op. cit.*, p. 9).

L'angle mort des effets de la mise en nombre des pratiques culturelles

Court état de l'art : le rendez-vous manqué des études sur la quantification et des *book studies*

- 7 Le corpus des travaux consacrés à l'étude de la quantification, dont nous ne pourrions fournir ici qu'un trop bref aperçu, est foisonnant, mais encore peu structuré ; comme le notent Berman et Hirschman, il est « encore loin de disposer de principes généraux ou d'un langage théorique unifié⁵ » (2018). Nous mobiliserons donc un ensemble diversifié de recherches francophones et anglophones issues d'une grande variété de disciplines, mais dont le trait commun est de s'intéresser « à la production et à la communication des nombres⁶ » (Espeland et Stevens, 2008) — ou, plus précisément, de proposer une « étude sociale [...] de ce que la mise en nombres *fait* plutôt que de ce qu'elle *est*⁷ » (Day, Lury et Wakeford, 2014). Notre intention est ici de souligner que, si vaste et si divers que soit ce corpus, il n'intègre que peu d'analyses de la « mise en données » des pratiques culturelles — et plus spécifiquement du sujet qui nous intéresse : la quantification des pratiques littéraires sur les plateformes numériques.

Penser scientifiquement les cultures de la quantification

- 8 Le champ d'études consacré à la quantification est largement dominé par les approches historiques — qu'elles proviennent d'historiens (Crosby, 2003 ; Porter, 2017), de philosophes (Hacking, 1990), de juristes (Supiot, 2015), d'anthropologues (Shore et Wright, 2024) ou de sociologues (Desrosières, 2008 ; Espeland et Stevens, 2008 ; Martin, 2020). Tous ces auteurs ont en commun d'analyser de façon souvent extrêmement détaillée « [l]e mouvement par lequel l'Occident est passé d'une perception qualitative à une perception quantitative de la réalité » (Crosby, *op. cit.*, p. 59). Sans nous attarder ici sur les origines historiques de l'essor de la quantification — qui tiennent à la conjonction d'éléments « tant bureaucratiques que scientifiques » (Porter, *op. cit.*, p. 33) —, nous en dirons simplement qu'elles reposent partiellement sur le développement du commerce ; comme l'écrit Theodore Porter, « il n'est guère possible d'imaginer une économie de marché et de commerce là où il n'y aurait ni prix ni mesures, ni par conséquent un usage considérable de la quantification » (*ibid.*, p. 36).
- 9 De fait, l'une des conséquences les plus extensivement documentées de la quantification porte sur la généralisation de ce que les anthropologues Cris Shore et Susan Wright ont nommé la « culture de l'audit » (*op. cit.*), qui se manifeste par l'immixtion de la « raison statistique » (Desrosières, 2010) dans la (quasi-)totalité des sphères de la vie politique et sociale — de la démocratie (Alonso et Starr, 1987 ; Mennicken et Salais, 2022 ; Rose, 1991) au marché du travail et au droit (Supiot, *op. cit.*), en passant par la santé (Ruckenstein et Schüll, 2017) ou la psychologie (Hacking, 2006 ; Rose, 1988), du monde universitaire (Espeland et Sauder, 2007) au football (Beer, 2015) ou aux réseaux sociaux (Grosser, *op. cit.*). Empruntant à Philip Mirowski le concept de « néolibéralisme de tous les jours⁸ », David Beer invite dès lors à considérer la façon dont les « systèmes de mesure et de compétition sont devenus une part de notre vie quotidienne, de nos imaginaires et de nos pratiques⁹ » (2015). Pour le dire autrement, nous sommes baignés dans des « écologies numériques¹⁰ » impliquant de « vivre dans et

avec les nombres¹¹ » (Day, Lury et Wakeford, *op. cit.*) — et ainsi obligés de subir le contrôle qu'ils imposent aux corps et aux subjectivités (Foucault, 2004 [1979] ; Citton, 2016). Ces « technologies de gouvernement » (Rose et Miller, 1992), incorporées en « techniques de soi » (Foucault, 2001[1984], p. 1358-1380), conduisent en effet les individus à adopter des comportements oscillant entre « managérialisation de soi », « optimisation », auto-contrôle (Dagiral, Licoppe *et al.*, 2019 ; Dalgarrondo et Fournier, 2019), surveillance (Whitson, 2013 ; Albrechtslund et Lauritsen, 2013) et « sousveillance » (Selke, 2016).

- 10 Les formes individuelles de quantification ont par ailleurs également fait l'objet d'un ensemble relativement important de travaux, souvent consacrés à l'expansion du *self-tracking* — pratique visant « à régulièrement surveiller, enregistrer et souvent mesurer certains aspects du comportement ou des fonctions corporelles d'un individu¹² » (Lupton, 2016, p. 2). L'auto-quantification a ainsi été analysée au travers d'exemples concrets aussi variés que le *tracking* du sommeil (De Cristofaro et Chiodo, 2023 ; Calvignac, 2021), les données physiologiques (Ruckenstein, 2014 ; Dagiral, Dessajan *et al.*, 2019), la sexualité (Lupton, 2015), le cycle menstruel (Thomé, 2023), les émotions (Pritz, 2016), le nombre de pas quotidiens (Pharabod, 2019), les pratiques sportives (Lopez-Gonzalez, 2014 ; Till, 2014 ; Whitson, *op. cit.*) — en somme, majoritairement à travers la question du corps.
- 11 L'expression « cultures de la quantification » que nous proposons ici vise à appréhender la façon dont les formes chiffrées — des indicateurs statistiques aux compteurs, en passant par les pourcentages, graphiques, métriques, notes et autres systèmes de points, etc. — se sont progressivement immiscées dans la quasi-totalité des sphères des sociétés occidentales contemporaines. Bien loin de n'être que des abstractions mathématiques, les nombres sont aujourd'hui mobilisés pour leurs propriétés sociales bien concrètes : leur efficacité communicationnelle et leur capacité à circuler largement dans des espaces socioculturels hétérogènes (Espeland et Stevens, 1998), la confiance qu'ils inspirent ainsi que leur aura d'objectivité et de scientificité (Porter, *op. cit.*), mais également leur capacité à influencer les comportements qu'ils prétendent décrire et à favoriser un rapport managérial au soi, plus productif, plus responsable et transparent (*accountable*) (Muller, 2018). Devenue omniprésente, ordinaire et in fine souvent invisible, la quantification altère le rapport que les individus entretiennent à leur environnement social et matériel, à leurs pratiques les plus quotidiennes et intimes, à leurs proches et à eux-mêmes. De fait, si les métriques ont la capacité de « fabriquer des individus¹³ » (Hacking, 2006) en créant de nouvelles catégories statistiques, elles ont également le pouvoir de les « intrastructurer », pour reprendre à notre compte un néologisme proposé par Yves Citton (2017, p. 138). Cette notion d'intra-structuration permet de penser la façon dont les médias numériques influencent, coconstruisent, altèrent les subjectivités de leurs utilisateurs. Le concept de « subjectivité », finalement rarement défini, sera ici entendu au sens foucauldien — un « mode de "subjectivation" » étant déterminé par « ce que doit être le sujet, à quelle condition il est soumis, quel statut il doit avoir, quelle position il doit occuper dans le réel ou dans l'imaginaire, pour devenir sujet légitime » d'un régime de connaissances (Foucault, 2001 [1984], p. 1451).
- 12 C'est d'ailleurs en nous inspirant des expressions « cultures algorithmiques¹⁴ » (Seyfert et Roberge, 2016) et « cultures du *self-tracking* » (Lupton, 2016) que nous proposons ici de parler de « cultures de la quantification ». Comme l'indique Deborah Lupton, le

terme cultures permet « de rendre compte de l'idée selon laquelle [...] les pratiques, significations, discours et technologies associées au *self-tracking* sont, de façon inhérente et inévitable, le produit de processus sociaux, culturels et politiques plus larges¹⁵ » (*ibid.*, p. 1). La formule nous semble en ce sens particulièrement pertinente pour rappeler que la mise en nombre des « formes culturelles du quotidien¹⁶ » (Beer, 2015, p. 11) ne relève pas seulement de pratiques individuelles et s'inscrit avant tout dans le cadre d'un processus historique, économique et politique de temps long et de grande ampleur. De fait, et si certains auteurs définissent la quantification de soi comme des « mesures [ayant] en commun d'être des notations chiffrées personnelles introduisant une forme de réflexivité ou de retour sur soi volontaire » (Pharabod, Nikolski et Granjon, 2013), il nous semble au contraire que, dans un monde structuré par la quantification, la mesure de soi ne saurait être présentée comme reposant exclusivement sur la volonté des individus.

- 13 Ce rapide panorama des analyses produites au sujet de l'(auto-)quantification n'avait pas d'autre objectif que de souligner en creux le surprenant manque d'attention porté aux effets spécifiques de la quantification culturelle. Dans la plupart des ouvrages s'intéressant aux processus de *datafication*, la culture n'est pas mentionnée ; au mieux, elle est traitée au détour d'un chapitre comme exemple parmi les autres de la mise en nombre des sociétés (Martin, *op. cit.*, p. 166 ; Cardon, 2015) ou selon une perspective historique s'intéressant aux statistiques culturelles produites par les institutions publiques (Dubois, 2015), mais négligeant d'analyser la présence sur Internet de formes de quantification culturelle « ordinaires » ou « infra-ordinaires » — au sens d'omniprésentes, banales (Stenger et Coutant, 2009 ; Jeanneret et Souchier, 2005).

Book studies et quantification

- 14 En parallèle de ces travaux, la recherche sur le monde du livre contemporain (que nous appellerons ici *book studies*) s'est, elle aussi, très peu intéressée aux conséquences de la quantification sur les lecteurs et leurs pratiques. Les études consacrées à la mise en nombre de la culture traitent généralement la question sous l'angle spécifique de la « quantification de la vie culturelle », c'est-à-dire en se concentrant sur l'« appareillage statistique [...] élaboré dans un cadre national par les établissements publics » (Donnat, 2020). Le champ littéraire (a fortiori francophone) dispose de fait d'une longue tradition de quantification par les institutions publiques, qui commanditent et supervisent à intervalles réguliers la production de rapports statistiques destinés à renseigner les pratiques culturelles des citoyens et à soutenir les politiques publiques (Böhm et Land, 2009 ; Dubois, 2012). La recherche s'est donc davantage intéressée à la façon de produire de meilleures statistiques ou de meilleures politiques publiques (MacDowall *et al.*, 2015) — ou à la façon dont la littérature a été intégrée à l'économie des plateformes numériques et à la « culture algorithmique » (Murray, 2018, 2021 ; Striphas, 2009, 2010) — qu'aux effets de la quantification des pratiques littéraires sur le rapport que les individus entretiennent à la lecture et à la littérature.
- 15 Une exception notable réside cependant dans le travail de Marine Siguier, et particulièrement son livre le plus récent, dont une petite partie traite de la mise en nombre de la lecture sur les réseaux sociaux et les plateformes littéraires. S'intéressant à des défis littéraires comme le « *challenge* à 1 000 » — consistant à « lire au moins 1000 pages (romans, BD, mangas, livre audio...) entre le vendredi 19h et le dimanche 23h59 »¹⁷ —, elle souligne que « [t]out en mobilisant l'idéologie d'une lecture gratuite

“pour le plaisir”, le *challenge* repose sur des enjeux productivistes » (2025, p. 205), ajoutant qu’ils reflètent « le succès d’une conception de la littérature comme pratique quantifiable » (*ibid.*, p. 207).

- 16 De la même façon, Oriane Deseilligny a analysé la « prescription comptable » à l’œuvre sur les plateformes qui, comme Amazon, organisent la recommandation de livres autour de « dispositifs d’évaluation fondés sur la quantification, exprimée par divers vocables tels que “palmarès”, “top”, “ranking”, “classement”, “taux”, “moyennes”, “notes”, etc. » (2014, p. 40). Elle observe alors très justement que « le caractère mathématique et statistique du système [...] fonctionne comme un argument *per se* » (*ibid.*, p. 41). Si elle parle ici de la recommandation de livres — un bon livre sur Amazon étant dès lors un livre ayant reçu un nombre suffisant de (bonnes) notes et de critiques —, ce constat s’applique également aux contenus littéraires sur les plateformes, où les nombres constituent autant d’arguments venant attester du niveau d’expertise, de popularité ou de productivité des utilisateurs. Il y a donc, conclut Deseilligny, une « prescription par la forme » aussi bien que par le fond — prescription de livres, mais également, ajoutons-nous, prescription d’attitudes à adopter, de scripts à suivre (Akrich, 1987).
- 17 Précisons cependant pour compléter ce panorama que le manque d’attention portée à la quantification des pratiques littéraires n’est pas exclusif aux chercheurs : la question du *tracking* des pratiques culturelles semble également très peu investie par les adeptes de l’auto-quantification eux-mêmes. Le site officiel du mouvement Quantified Self recense ainsi les centaines de présentations (appelées *Show and Tell*) au cours desquelles les membres du mouvement ont présenté leurs projets d’auto-quantification. Si les centaines d’exposés publiés depuis 2007 portaient sur des questions aussi hétérogènes que le cholestérol, le temps d’écran, l’acné, les émissions carbone, la consommation de bière, la ponctualité ou la maladie de Parkinson, nous n’avons relevé que deux expérimentations relevant véritablement de la quantification littéraire¹⁸. Le « *lifelogging* » (Selke, *op. cit.*) accorde donc une place pour le moins anecdotique à la mise en nombre des pratiques littéraires — qui tranche avec l’omniprésence de la quantification de ces activités sur Internet.
- 18 Ce constat nous amène à penser, comme nous tenterons de le démontrer plus bas, que l’(auto-)quantification des pratiques littéraires est avant tout une pratique imposée par les plateformes et qu’elle est investie par les individus sur un mode beaucoup moins volontaire que ne le sont traditionnellement les pratiques de *self-tracking*. Il est d’ailleurs possible de supposer que la quantification des pratiques littéraires produit d’autant plus d’effets qu’elle est insidieuse, peu conscientisée, peu étudiée.

L’incontournable référence foucauldienne

- 19 Finalement, et comme l’écrivent Éric Dagiral et ses coauteurs dans leur revue de la littérature sur le *quantified self*, « Michel Foucault constitue une référence très fréquente, voire obligée, dans la littérature sur le *self-tracking* » (2019, p. 31). C’est, ajoutons-nous, plus globalement un passage obligé de toutes les études portant sur la quantification. Les usages des concepts foucauldien, si désordonnés, hâtifs ou nébuleux qu’ils soient parfois, apparaissent en effet incontestablement utiles pour éclairer la mise en nombre des sociétés et le « projet néolibéral » (Gane, 2014) qui la sous-tend souvent. Theodore Porter note de fait que « [l]’intérêt contemporain pour les normes et les indicateurs, lié à ce qu’on appelle souvent la “gouvernance néolibérale”, a aussi

encouragé un souci des chiffres » (*op. cit.*, p. XV) — ce que Foucault étudie notamment dans *Naissance de la biopolitique* (2004 [1979]), où il tente d'explicitier le fonctionnement du (nouveau) système « économico-institutionnel » que constitue le néolibéralisme (Laval, 2014). L'usage de ce terme appelle bien sûr des précautions — notamment parce que « le néolibéralisme n'existe sous aucune forme pure¹⁹ » (Beer, 2016, p. 20), que sa variabilité historique et géographique en fait parfois un « signifiant trop lâche²⁰ » (Brown, 2017, p. 20) et que « [l]e concept de néolibéralisme peut constituer un point de départ » pour l'analyse, « mais il ne devrait certainement pas être une conclusion en soi »²¹ (Peck, 2010, p. 13). Il reste cependant utile pour penser la place accordée à la concurrence et à l'entreprise comme modèles de société. Foucault postule en effet que la « nature spécifique du néolibéralisme » tiendrait à sa tendance à « faire de l'entreprise un modèle général », dont la vocation serait de s'étendre bien au-delà de la simple sphère économique, jusque dans les relations interpersonnelles, intra-familiales, et même dans la relation que les individus entretiennent à eux-mêmes (Laval, *op. cit.*).

- 20 Cette pensée est centrale pour expliquer l'immixtion progressive des rationalités néolibérales et de la quantification dans les comportements individuels. Christian Laval écrit ainsi dans une optique foucaldienne que « [l]e maître mot du nouveau rapport de soi à soi c'est le dépassement de soi-même, l'ultrasubjectivation, une figure inédite de la subjectivité ordonnée à la logique de l'entreprise » (*ibid.*, p. 156). Cette « agentivité néolibérale²² » (Gershon, 2011) conduirait à penser « le soi comme entreprise²³ » (McNay, 2009). Comme Foucault le formule lui-même, au sein d'un régime néolibéral,

[i]l faut que la vie même de l'individu, avec par exemple son rapport à sa propriété privée, son rapport à sa famille, à son ménage, son rapport à ses assurances, son rapport à sa retraite, fasse de lui comme une sorte d'entreprise permanente et d'entreprise multiple (2004 [1979], p. 247).

- 21 En d'autres termes, et selon la formule de Dardot et Laval, « “entreprise” est le nom que l'on doit donner au gouvernement de soi à l'âge néolibéral » (2010, p. 409) — ce qui, nous le montrerons, se retrouve (de façon toutefois nuancée) dans les comportements des lecteurs imprégnés par les cultures de la quantification.
- 22 La référence foucaldienne est par ailleurs ici d'autant plus incontournable que les pratiques littéraires et l'auto-quantification ont en commun de constituer des « technologies de soi », c'est-à-dire des « pratiques réfléchies et volontaires par lesquelles les hommes non seulement se fixent des règles de conduite, mais cherchent à se transformer eux-mêmes, à se modifier » (Foucault, 2001 [1983], p. 1364). Comme le note Aude Volpilhac, Foucault souligne en effet que « la lecture fait partie, notamment chez les Stoïciens, les Sceptiques et les Épicuriens, des exercices de pratique de soi qui contribuent à la constitution du sujet » (2015, p. 36). En d'autres termes, les pratiques littéraires participent souvent d'une « culture de soi » (ou d'un « souci de soi ») caractérisée par « le fait que l'art de l'existence [...] [se] trouve dominé par le principe qu'il faut “prendre soin de soi-même” » (Foucault, 2014 [1984], p. 60-61). Ce rapport à soi particulier présente dès lors bien des points communs avec les subjectivités néolibérales contemporaines, qui invitent « à se manager de façon réflexive comme si le soi était une entreprise²⁴ » (Gershon, *op. cit.*, p. 546).
- 23 Dans ce cadre, les pratiques littéraires peuvent être perçues comme une « discipline » à laquelle les lecteurs se plient « dans une optique de progression » (Barth-Rabot, *op. cit.*,

p. 95). Il s'agit dès lors de considérer la dimension ascétique de la lecture, au travers de laquelle le sujet — en acceptant de mettre entre parenthèses son corps et ses préoccupations extérieures — fournit la preuve de son désir de s'améliorer (*ibid.* ; Volpillac, *op. cit.*). Cette conception de la lecture comme ascèse n'est pas nouvelle (Mauger et Poliak, 1998 ; Bourdieu, 2015 [1992] ; Lahire, 2006), mais prend certainement un sens nouveau à l'ère où les pratiques littéraires sont concurrencées par des activités requérant un moindre niveau d'attention (Barth-Rabot, *op. cit.*). Plus que jamais, la lecture est donc présentée comme une pratique exigeante, mais gratifiante en ce qu'elle serait « le prix à payer » pour s'améliorer et « avoir accès à la vérité » (Volpillac, *op. cit.*, p. 37).

Méthodologie et présentation du terrain

- ²⁴ Afin d'étudier les effets de la quantification des pratiques littéraires sur les lecteurs, sur une période s'étendant de mai 2024 à février 2025, nous avons donc conduit 31 entretiens semi-dirigés avec des lecteurs²⁵ francophones (24 femmes et 7 hommes, 19 Français et 12 Belges²⁶), ayant pour trait commun d'utiliser des plateformes numériques pour consulter ou produire du contenu lié aux livres. Parmi eux, 17 étaient actifs sur Babelio, 6 sur Livraddict et 11 sur Instagram²⁷ — tous étant inscrits sur ces plateformes depuis plus d'une année. Nous indiquerons à chaque citation d'entretien leur plateforme de prédilection entre parenthèses, mais soulignons tout de même que la majorité de nos enquêtés ne se limitent pas à un seul réseau et mobilisent plusieurs sources numériques d'information sur les livres, de Facebook à Twitch, en passant par Discord ou YouTube.
- ²⁵ Nous avons en partie sélectionné les personnes interrogées sur le critère de l'intensité de leurs pratiques littéraires : ainsi, et même si certains en ont douté, nos enquêtés sont tous de « grands lecteurs²⁸ », déclarant lire entre 30 et 300 livres par an — la grande majorité d'entre eux se situant entre 50 et 100 lectures. 24 d'entre eux participent également à des *challenges* de lecture (formes de jeux littéraires qui invitent à « valider » des lectures suivant une ou plusieurs contraintes, sur une période souvent limitée, parfois illimitée) et 10 sont impliqués dans l'organisation d'un ou plusieurs de ces défis littéraires. Sept lecteurs exercent finalement une activité de créateurs de contenu littéraire sur Instagram auprès de communautés relativement restreintes à l'échelle de la plateforme (leurs comptes regroupaient entre 1 000 et 8 000 abonnés), mais significatives pour des contenus littéraires et suffisantes pour entretenir des relations avec les maisons d'édition, recevoir des livres dans le cadre d'envoi de « services presse » et être affilié à des programmes publics de promotion de la lecture²⁹.
- ²⁶ Parce que la circulation des artefacts quantifiés ne s'arrête pas à la frontière des plateformes et que toutes n'attirent pas le même public, nous avons donc choisi de nous entretenir avec des lecteurs inscrits sur deux types de plateformes : un réseau social, Instagram, et deux réseaux littéraires francophones : Babelio, premier acteur de recommandation littéraire revendiquant plus de deux millions d'inscrits, et Livraddict, un site au succès plus confidentiel, mais ayant constitué une solide base d'adeptes depuis 2009 et reposant sur le bénévolat. Malgré ces différences de taille, les raisons qui ont motivé nos enquêtés à rejoindre une plateforme littéraire sont identiques d'un site à l'autre : ils cherchaient à se constituer une bibliothèque virtuelle, à obtenir des recommandations de lecture, et à échanger avec d'autres lecteurs — notamment par

l'intermédiaire des défis organisés par des membres, qu'ils ont souvent découverts par hasard, en parcourant le site. Babelio et Livraddict présentent en outre des interfaces comparables, particulièrement dans les fonctionnalités de quantification qu'elles offrent à leurs utilisateurs, qui sont exposés à une profusion de métriques et de graphiques censés décrire leurs pratiques de lecture : statistiques sur leurs genres littéraires, auteurs, éditeurs, nationalités d'auteurs ou périodes historiques, etc., préférés, mais également nombre de livres lus au cours de l'année ou des années précédentes, nombre de lectures abandonnées ou en cours, pourcentage d'avancement dans les lectures du moment, taille de leur bibliothèque virtuelle ou de la « pile à lire »... À cela s'ajoutent encore le décompte des critiques rédigées, des mentions « j'aime » reçues, des abonnés et des abonnements — la liste pourrait encore être prolongée.

- 27 Outre ces entretiens menés avec les utilisateurs de ces plateformes, nous avons également choisi d'interroger des créateurs de contenus littéraires sur Instagram — où bilans chiffrés, statistiques, notes et compteurs de lecture sont aussi très populaires. Il existe d'ailleurs une grande porosité entre les plateformes littéraires et Bookstagram³⁰ : un certain nombre de nos enquêtés (notamment les plus jeunes) disposaient de comptes sur plusieurs réseaux et utilisaient les données produites par les plateformes littéraires pour produire du contenu sur Instagram.
- 28 Il nous paraît ainsi important de préciser que les plateformes n'ont pas le monopole de la production de métriques. Il existe un certain nombre de biais par lesquels les lecteurs produisent eux-mêmes des indicateurs statistiques : sur les réseaux sociaux, bien sûr, où les créateurs de contenu réalisent leurs propres visuels, mais également sur les plateformes littéraires. Sur Babelio et Livraddict, les organisateurs de *challenges* ont mis en place leurs propres systèmes de quantification en créant des fichiers Google Sheets (équivalents à des fichiers Excel partagés en ligne), consultables par tous et synthétisant grâce à des pourcentages et à des systèmes de paliers l'avancée de tous les participants à un *challenge*³¹.
- 29 Nous étudierons donc ici de façon indifférenciée l'ensemble des artefacts qui visent à mettre en nombre les pratiques littéraires, entendues « comme un ensemble d'actions et de discours » (Barth-Rabot, *op. cit.*, p. 12).). Compteurs, graphiques, pourcentages, ratios, classements... Nous tenterons d'analyser les effets des « petites formes » (Candel *et al.*, 2012) quantifiées — qui structurent l'interface des plateformes littéraires et parsèment le contenu des influenceurs — sur les activités qu'elles prétendent décrire. Le choix d'étudier ensemble des artefacts quantifiés aussi divers et produits sur des plateformes aussi différentes que Babelio, Livraddict et Instagram peut paraître surprenant. Il nous semble toutefois qu'ils participent d'une même « raison statistique » (Desrosières, 2010), d'une même conception chiffrée des pratiques littéraires propagée d'un espace numérique à l'autre — bref, des mêmes cultures de la quantification. Ainsi, et à la suite d'Olivier Martin,

nous défendons l'idée qu'il est possible de penser, dans leur globalité, tous les chiffres qui nous entourent, quels que soient leur nature apparente et le rôle qu'on leur accorde spontanément. Ils ne sont pas tous identiques, mais ils ont un socle commun, des raisons d'être partagées et des effets similaires (*op. cit.*, p. 3).
- 30 Notons le statut très particulier des algorithmes et des notes : les uns, invisibles, se nourrissent de la quantification mais ne la matérialisent pas, tandis que les autres ne

visent pas tant à mettre en nombre des pratiques que des goûts ; aussi les laisserons-nous majoritairement de côté ici.

- 31 Par ailleurs, et malgré la diversité des profils socioculturels de nos enquêtés, aucun lien univoque ne semble pouvoir être tracé entre leur rapport aux statistiques et leur âge, leur classe sociale ou l'emploi qu'ils occupent (ou occupaient). Il apparaît seulement que l'intérêt pour la quantification est légèrement moins explicite chez les personnes ayant depuis longtemps une habitude de lecture assidue, qui étaient aussi souvent les lecteurs les plus âgés et les retraités. Il nous faut tout de même souligner que ces enquêtés participaient souvent à des défis de lecture impliquant des statistiques, paliers et autres systèmes de points, ce qui suggère tout de même un attrait implicite pour la mise en nombre. Ce constat tendrait à suggérer que la quantification, quand elle est explicitement recherchée, peut être utilisée comme source de motivation pour des lecteurs qui n'ont pas toujours lu assidûment (voire qui, pendant une partie de leur vie, n'ont pas lu du tout) afin de maintenir un rythme de lecture qui ne leur est pas (ou plus) coutumier.
- 32 Toutefois, et comme le note Cécile Barth-Rabot, « la valeur symbolique de la lecture et la dimension identitaire d'une pratique qui est souvent aussi un positionnement, ont des implications méthodologiques » (*op. cit.*, p. 21). Nous considérerons donc, comme elle le suggère, les discours de nos enquêtés « comme étant conjointement un témoignage sur leur pratique et sur leur système de valeurs et, *de facto*, un positionnement dans un espace social où lire ou ne pas lire n'est jamais anodin » (*ibid.*).

Pour une analyse critique des effets de la quantification des pratiques littéraires

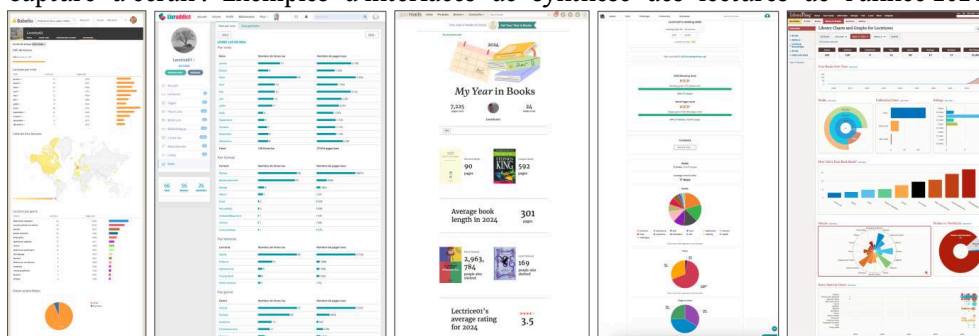
Standardisation et influence de la quantification littéraire sur les plateformes

Omniprésence, esthétisation et viralité des nombres

- 33 Les plateformes littéraires sont constellées d'artefacts visuels quantifiés qui, du compteur aux graphiques, agissent comme systèmes de recommandation, c'est-à-dire comme des « outils et des techniques fournissant des suggestions d'items susceptibles d'être utiles à un utilisateur³² » (Ricci, Rokah et Shapira, 2022, p. 1). Placés au cœur des interfaces, ils permettent de hiérarchiser les livres, les membres, les auteurs — associant implicitement la qualité et le niveau d'expertise aux nombres (de leurs lectures, de *likes*, d'abonnés, d'étoiles, de pourcentages).
- 34 Ces artefacts quantifiés se diffusent sur le Web sans qu'il soit possible de démêler avec certitude leur point d'origine : s'ils sont présents sur l'ensemble des principales plateformes littéraires — des leaders anglophones (Goodreads, LibraryThing, The Storygraph) à leurs concurrentes francophones (Babelio, Livraddict Booknode, SensCritique), hispanophones (Lecturalia), germanophones (LovelyBooks) —, ils circulent également dans les contenus des influenceurs, les discours des utilisateurs, les outils créés par des lecteurs³³ ou les campagnes pour la lecture³⁴.
- 35 C'est d'ailleurs l'une des propriétés centrales des nombres que de circuler plus facilement que les formes textuelles — devenant une monnaie d'échange des

conversations, une marque d'autorité, du contenu pour les réseaux : « Les nombres sont faciles à déloger de leurs contextes locaux pour être réinsérés dans des contextes plus éloignés³⁵ » (Espeland et Sauder, 2007). L'un des exemples les plus probants de cette diffusion des artefacts visuels de quantification est sûrement celui du format « bilan de lecture ». Comme le montre la série de captures d'écran exposée ci-dessous, les principales plateformes littéraires francophones et anglophones ont toutes adopté des codes similaires pour synthétiser les pratiques littéraires de leurs utilisateurs, créant des pages saturées de nombres et de graphiques colorés, souvent calquées sur le format du bilan annuel de Spotify (le *wrapped*), conçu pour être repartagé massivement sur les réseaux sociaux (Annabell et Rasmussen, 2024), comme illustré par la seconde paire de captures d'écran.

Capture d'écran. Exemples d'interfaces de synthèse des lectures de l'année 2024.



Légende : De gauche à droite : Babelio, Livraddict, Goodreads, TheStorygraph, LibraryThing. Sources : Captures d'écran réalisées par l'autrice en février 2025.

36 Captures d'écran et 3. Circulation des formes quantifiées sur les réseaux sociaux



- 37 Sources : (gauche) Vidéo YouTube de Destiny Sidwell, 1^{er} janvier 2024 ; (droite) Publication Instagram de pandorasboxofbooks, 6 décembre 2024.
- 38 Or, « la provenance d'un nombre et sa destination influencent ce que ce nombre est et ce qu'il fait³⁶ » (Day, Lury et Wakeford, *op. cit.*) : parce qu'elles sont souvent produites à destination des réseaux sociaux, ces formes quantifiées sont conçues pour être virales, et donc esthétiques. Plusieurs enquêtés ont ainsi souligné le plaisir visuel qu'ils éprouvent à consulter leurs statistiques : « J'aime bien avoir des jolis graphiques en couleurs et j'ai plein de chiffres écrits [...] que j'aime bien avoir ordonnés avec des barres » (Alice, 23 ans, étudiante, Livraddict). Cet attrait esthétique a également pour effet d'intensifier certaines pratiques littéraires. Pour ne prendre qu'un seul exemple,

Babelio et Livraddict proposent une carte du monde qui se colore en fonction de la nationalité des auteurs lus — ce qui, pour plusieurs enquêtés, s’est transformé en jeu : celui de terminer un ouvrage écrit par un écrivain de chaque pays du monde. « J’aime beaucoup la carte. La carte du monde dans les statistiques de lecture. [...] Du coup, j’essaie de la remplir ! » (Élise, 38 ans, ingénieure, Babelio). Dès lors, et de la même façon que Grosser se demandait au sujet de Facebook si nous « ajouterions [...] autant d’amis si nous n’étions pas constamment confrontés à leur nombre³⁷ » (*op. cit.*), nous nous interrogeons : les lecteurs liraient-ils autant si, de leur fil d’actualité Instagram à leur page personnelle Babelio, ils n’étaient pas exposés à des histogrammes décrivant des performances de lecture, des pourcentages d’avancement dans leurs objectifs, des photos de piles de livres, des vidéos « bilan des lectures du mois », etc. ?

Transparence imposée et banalisation des pratiques de quantification

- 39 Cette interrogation est d’autant plus légitime que, contrairement à Facebook, les plateformes littéraires n’offrent aucune possibilité de restreindre l’accès à ses métriques personnelles : non seulement la quantification individuelle des pratiques littéraires n’est pas optionnelle (tous les utilisateurs de Babelio, Livraddict, Goodreads, etc., disposent d’une page « Mes statistiques »), mais elle est aussi nécessairement publique (tous les internautes peuvent accéder à cette page). Pour Miriam, cette publicité — qu’elle apparente à du voyeurisme, malgré le pseudonymat que permet le site — tranche avec l’intimité souvent associée aux pratiques littéraires : « la première fois que j’avais vu [la page « Mes statistiques »] sur un autre profil, ça m’a fait hyper bizarre. [...] Parce que c’est tellement détaillé que... [...] je me suis dit “Waouh, ça fait hyper bizarre, c’est un peu voyeuriste !” » Elle précise :

Les statistiques, on ne les a ni créées ni... Elles sont faites par la plateforme, elles sont dans une rubrique et elles concernent nos habitudes de lecture. Et nous-mêmes, si on ne fait pas toutes ces additions-là, on ne les a pas en tête, on ne sait pas qu’on a lu tel nombre [de livres]. [...] Donc je trouve ça beaucoup plus intime que la critique qu’on a écrite tout en sachant [qu’elle serait lue] (33 ans, aide-soignante, Babelio).

- 40 C’est toute la tension, pour reprendre les termes de Miriam, entre intimité et voyeurisme, entre dévoilement volontaire et exposition contrainte qui pose la question de la « sousveillance », c’est-à-dire la surveillance des utilisateurs entre eux (Lupton, 2016).
- 41 Comme le note Lupton en reprenant à son compte la pensée de Donna Haraway, « [l]es assemblages homme-données [...] produisent également des effets matériels sur les humains de manière réursive » (2020, p. 17). Cette idée d’un « bouclage réursif » — caractérisé, en termes électroniques, par le fait que le produit (*output*) d’un circuit ait une incidence sur son entrée (*input*) (Citton, 2017, p. 67) — est centrale pour comprendre le rapport que les lecteurs entretiennent aux artefacts quantifiés. C’est parce qu’ils sont quotidiennement exposés à la mise en nombre des activités de lecture et parce que « ce qui circule d’un point A vers un point B ne se contente pas d’altérer ce dernier, mais fraie la voie, par sa circulation même, aux flux qui le suivront » (*ibid.*, p. 117), que les lecteurs interrogés ici ont tous accepté l’idée, malgré leurs réticences initiales, leurs critiques, voire leur ironie, que leurs pratiques littéraires seraient traduites en statistiques, en graphiques, en classements et que toutes ces formes

visuelles (sur lesquelles ils exercent un contrôle très marginal) seraient mises en scène aux yeux de tous.

- 42 De fait, aucun de nos enquêtés n'ignorait la présence de ces données chiffrées — même si rares étaient ceux qui déclaraient s'être inscrits sur une plateforme dans le but d'obtenir des statistiques sur leurs pratiques littéraires. La majorité du temps, les raisons invoquées étaient le désir de garder la trace de leurs lectures, d'obtenir des recommandations de livres ou d'interagir avec d'autres lecteurs. Ainsi, et si une partie importante des lecteurs interrogés n'effectuait pas fréquemment la démarche volontaire de consulter les métriques produites, ils y faisaient tout de même face, ne serait-ce qu'en raison du design des interfaces : « Je suis obligée de le regarder parce que c'est tout en haut de la page d'accueil, donc je le vois » (Agathe, 29 ans, éducatrice spécialisée, Livraddict). En somme, « une fois les utilisateurs engagés sur ces plateformes, un self-tracking à la fois volontaire et involontaire se met en place³⁸ » (Lupton, 2016, p. 74).
- 43 La plupart des personnes interrogées interagissent dès lors avec ces données sur un mode beaucoup moins actif et délibéré que ne peuvent le faire des *self-trackers* portés sur le sport ou la santé. Pour beaucoup, les statistiques sont mobilisées parce qu'elles sont là : « C'est Babelio qui le fait pour moi ! Moi je compte pas. Il y a un compteur sur Babelio, on enclenche le compteur en début d'année. Sinon je compterais pas, sincèrement » (Anne, 46 ans, professeure de français en collège, Babelio). Globalement, nos enquêtés entretiennent un rapport ambigu aux données : souvent indifférents ou perplexes face à la profusion des statistiques mises à disposition par les plateformes, ils en jouent néanmoins le jeu. La quasi-totalité d'entre eux s'est fixé un objectif de lecture (général, ou dans le cadre d'un *challenge* de lecture), souvent maintenu dans le temps par habitude :
- En fait, ils te demandent sur le côté. Ils te mettent « Est-ce que vous avez fixé votre défi annuel ? », et tout. Donc j'ai regardé, je me suis dit : « Oh bah c'est sympa », et puis ben je me suis lancée comme ça, en fait. [...] Depuis, je le fais tous les ans. (Audrey, 31 ans, fonctionnaire territoriale, Babelio).

L'incarnation graphique d'un modèle économique ?

- 44 La quantification est donc un choix d'interface aux effets concrets. Si notre objectif n'est pas ici de critiquer les artefacts quantifiés « de l'extérieur, en en faisant les reflets des intérêts de leurs concepteurs », mais plutôt de tenter de « comprendre de l'intérieur la manière dont ils produisent des effets (plus ou moins critiquables) sur nos sociétés » (Cardon, *op. cit.*, p. 15-16), il nous semble tout de même important de souligner les intérêts économiques directs que les plateformes peuvent tirer de la quantification. Parce qu'elle favorise l'intensification des pratiques de lecture, la mise en nombre favorise également une hausse du trafic sur le site, du nombre de publicités visionnées et de données commercialisables générées. Or, le modèle économique de Babelio, par exemple, repose essentiellement sur la vente d'encarts publicitaires aux annonceurs, de données bibliographiques aux bibliothèques et d'informations sur le lectorat aux maisons d'édition — autant d'éléments dépendant de la taille et de l'implication de l'audience. Le cas de Goodreads, la première plateforme littéraire mondiale et concurrent anglophone de Babelio, est peut-être encore plus évocateur, puisque la plateforme est possédée par Amazon... qui est à l'origine d'une part importante (sinon majoritaire) des ventes de livres aux États-Unis et dans le monde.

Ces intérêts financiers ont nécessairement une incidence sur les choix de design des plateformes et sur la décision d'accorder à la quantification une place si centrale.

- 45 En revanche, et sans omettre totalement ces intérêts économiques, il s'agit ici de ne pas en tirer de relation de cause à effet trop systématique — notamment parce que la plupart des entreprises n'ont pas toujours une vision stratégique aussi finement ciselée que ne le laissent penser leurs discours, voire parce que leurs pratiques peuvent être marquées par une forme d'opportunisme (multiplier les indicateurs statistiques parce qu'elles le peuvent, qu'elles disposent des données ou simplement pour imiter leurs concurrents). Notre objectif est donc de considérer les pratiques de quantification non pas du point de vue de leurs raisons d'être supposées, mais à partir des effets concrets, matériels qu'elles ont sur les utilisateurs de plateformes littéraires, sur leurs façons de lire, sur le rapport qu'ils entretiennent aux livres et à la littérature.

Définir les lectures qui comptent

- 46 La quantification implique de réunir en un chiffre unique des éléments hétérogènes (Supiot, *op. cit.*, p. 197). Comme l'écrit Alain Desrosières, « quantifier, c'est convenir puis mesurer ». En d'autres termes, quantifier nécessite de définir des « conventions d'équivalence » — qui sont autant « de négociations, de compromis, de traductions, d'inscriptions, de codages, de procédures codifiées et répliquables, et de calculs » — précédant la mesure, qui n'est dès lors qu'une « mise en œuvre réglée de ces conventions » (2008, p. 10-11). Quantifier les pratiques littéraires implique donc d'abord de formaliser *ce que lire veut dire*. Les plateformes véhiculent ainsi une conception très spécifique des formes de lecture *qui comptent* et qui suppose d'avoir (1) lu un livre (2) pour la première fois, (3) dont la lecture a été enregistrée.

1. Un livre lu. Sur les plateformes étudiées, le seul moyen d'ajouter un livre à ses statistiques de lecture (de le *faire compter*) est de l'indiquer comme « lu ». Or, les compteurs imposent une conception binaire de la lecture (lu/non lu) qui ne tolère aucune nuance : quid des livres feuilletés ? Des livres abandonnés³⁹ ? Les plateformes ne fournissant pas de définition précise de ce que « lu » veut dire, chaque lecteur est dès lors implicitement chargé de le définir lui-même. Bien souvent, les personnes que nous avons interrogées ne s'autorisaient à compter une lecture comme « terminée » que si elles avaient lu chaque page de l'ouvrage, ce que déplore Valérie (50 ans, conseillère pédagogique) : « Je pense qu'il y a une confusion dans la tête des gens. Pour eux, lire c'est [...] finir un livre. [...] Ils voient ça comme un acte fini », ajoutant : « je pense qu'ils passent à côté » du plaisir apporté par le fait de simplement « consulter » ou « de revenir vers le livre, de l'abandonner, mais d'y retourner ». Cette exigence d'exhaustivité de la lecture a sûrement partie liée au respect de l'œuvre et à la « fétichisation de la main de l'auteur » (Chartier, 2015, p. 62) qui sont régulièrement revenus dans les discours et qui impliquent un certain sens de l'ascèse : « Quand je commence un livre, je le finis, même si je l'aime pas. C'est une espèce de respect par rapport à l'auteur, si tu veux. L'auteur a pris la peine d'écrire ce livre, donc moi je vais prendre la peine de le lire » (Béatrice, 56 ans, auto-entrepreneuse, Babelio). Cependant, et parce que faire croître ses statistiques est souvent perçu comme désirable, coexistent également des acceptions plus lâches du terme lu. Sur les réseaux sociaux, les influenceurs présentant leurs « bilans » de lecture semblent se sentir fréquemment obligés de formaliser leurs règles de comptage — à l'instar de la BookTubeuse américaine Destiny Sidwell, qui se justifiait au sujet de ses statistiques de lecture de 2023 : « Mes lectures, mes règles : si je lis plus de 50 % d'un livre, a fortiori 60 %, et que je ne le termine pas, je le compte quand même comme "lu", parce que j'ai passé beaucoup de temps à lire le livre. Et puisque ce sont mes lectures personnelles, je le

compte⁴⁰. » Cela se retrouve également sur certains défis littéraires : « Un abandon compte comme une lecture si tu as lu minimum 25 % du livre⁴¹. » En somme, compter un livre comme lu implique surtout de justifier du sérieux et des efforts mobilisés lors de sa lecture.

2. Un livre *lu pour la première fois*. Sur les plateformes littéraires, la relecture — si valorisée qu'elle soit socialement (Barth-Rabot, *op. cit.*) — est rarement prise en compte : il est souvent impossible d'ajouter plusieurs fois un même livre à sa bibliothèque virtuelle⁴². Si certaines plateformes disposent d'une option permettant d'indiquer plus d'une date de lecture pour un même ouvrage, il n'est pas permis de le noter deux fois (ce qui pourrait permettre de matérialiser l'évolution des goûts). De la même façon, les *challenges* littéraires (du moins la grande majorité d'entre eux qui, comme sur Babelio ou Livraddict, demandent de solliciter l'approbation de l'organisateur pour ratifier les lectures) ne comptabilisent que les lectures récentes et n'autorisent pas de valider plusieurs consignes avec un même livre. À bien des égards, un livre relu est donc un livre qui *ne compte pas*.
3. Un livre dont la lecture a été *enregistrée*. Lise, comme une grande partie de nos enquêtés, a ritualisé le moment où elle achève ses lectures : « La première chose que je vais faire, c'est de noter que j'ai fini dans mon fichier Excel » (38 ans, assistante dans une association, Instagram). Le fait de terminer un livre est, pour les lecteurs interrogés, souvent étroitement associé à l'action de l'enregistrer, c'est-à-dire de le consigner dans un registre, qui peut être un fichier sur son ordinateur, un carnet papier ou une plateforme numérique. Danièle explique par exemple : « J'ai une routine. Donc, quand j'ai fini un livre, je m'installe, j'écris ma chronique sur Babelio et puis je vais dans mon dossier "*Challenges*". Et là je regarde tous mes Excel en me disant : "bon, est-ce que ce bouquin correspond à tel *challenge*, tel *challenge* ?" J'ouvre tous les Excel où il peut correspondre et puis... Je m'y mets, quoi » (49 ans, directrice de projets dans une structure privée à mission sociale, Babelio). Cette façon de répertorier systématiquement ses lectures tient sûrement en partie à la peur de les oublier — qui est l'une des motivations principales invoquées par nos enquêtés pour expliquer leur inscription sur les plateformes —, mais également au désir de garder ses statistiques à jour. Comme le note Lupton, la visualisation a pour effet de « geler » des données « vivantes » (*lively*), « liquides » (2016, p. 45), et les plateformes, soucieuses de favoriser une production de données en flux tendu, ont donc à cœur d'offrir des fonctionnalités permettant de réaliser des enregistrements intermédiaires. Gaëlle (25 ans, chargée de projet dans une collectivité territoriale, Goodreads), par exemple, a cessé d'utiliser des marque-pages et reporte directement sur Goodreads le numéro de la page où elle suspend sa lecture, qui calcule en retour son pourcentage d'avancement dans le livre — ce qui fait de l'enregistrement une étape cruciale pour *faire compter* sa lecture.

Mise en nombre et intensification des pratiques littéraires

« J'essaie d'augmenter chaque année » : inflationnisme littéraire en contexte quantifié

- 47 Si la quantification induit une conception restrictive de *ce que lire veut dire*, elle véhicule également une injonction à lire *plus*. Reprenant à son compte la théorie de l'acteur-réseau de Bruno Latour pour analyser l'influence des nombres sur Facebook, Grosser avance en effet que les métriques disposent de leurs propres capacités agentives et qu'« il est utile de penser [à elles] comme ayant des intentions intrinsèques⁴³ » (*op. cit.*). Les métriques, soutient-il, ont donc une forme de volonté propre :
- 48 Je défends l'idée que les métriques ne sont pas plus immunisées que nous face au désir d'accumulation — qu'en tant qu'instrument du capital, la métrique est tout aussi

sensible au fétichisme capitaliste de la croissance. Les métriques veulent « s'empiler », devenir plus grandes, croître toujours plus ^{(ibid.)⁴⁴}.

- 49 Les métriques « veulent plus » — un plus indéfini et intransitif qui se retrouve dans le cas de la quantification littéraire. Une part considérable des lecteurs interrogés ont ainsi décrit la logique inflationniste induite par les plateformes, qui les poussait à revoir à la hausse leurs objectifs d'une période à l'autre : « J'essaie d'augmenter par rapport à chaque année. [...] Je me mets cinq de plus en fait tous les ans » (Audrey).
- 50 Cette logique inflationniste est également particulièrement visible sur les réseaux sociaux, où les livres sont appréhendés en masse — achetés en grand nombre⁴⁵, critiqués en série (comme dans cette vidéo, représentative à plus d'un titre, qui propose la notation et la critique de 148 livres en 20 minutes⁴⁶), ils font également l'objet de défis de lecture toujours plus impressionnants (30 livres en 30 jours⁴⁷, 500 livres en un an⁴⁸, etc.). Mentionnons ainsi la croissance exponentielle des *book hauls*, ces vidéos extrêmement populaires au sein des communautés de lecteurs au cours desquelles les influenceurs présentent leurs dernières acquisitions et qui prennent progressivement des dimensions vertigineuses : 600 euros dépensés dans des livres en 2 mois⁴⁹, ou 20, 50⁵⁰, 100⁵¹ ouvrages acquis en quelques semaines.
- 51 Les habitudes d'achat et d'accumulation de nos enquêtés — qu'ils qualifient fréquemment de « compulsifs » — reflètent également cette inflation littéraire : la « pile à lire⁵² » de Sylvie a été décuplée depuis le confinement, passant de 40 à 400 livres ; celle de Léa dépasse les 500 titres non lus, quand celle de Florence (60 ans, employée dans le secteur bancaire, Instagram) a grossi « de façon exponentielle », jusqu'à atteindre 700 ouvrages ; Danièle déclare quant à elle posséder chez elle plus de 1 800 livres à lire, tandis que Christian ajoute chaque année de 150 à 200 nouveaux ouvrages aux milliers qu'il possède déjà : « Je dirais que ça dépasse facilement les 5 000, mais si ça se trouve, c'est 4 000 ou 8 000 » (73 ans, retraité, anciennement notaire, Babelio). La pile à lire (ou « PAL ») est un motif très récurrent des contenus littéraires sur Internet et son évocation entraîne souvent le même constat : elle grandit inexorablement.

Tenir le rythme : lecture sous contraintes temporelles

- 52 Ces logiques d'inflation et ces injonctions à la lecture se heurtent néanmoins à la réalité matérielle des lecteurs, qui ont très souvent évoqué lors de nos entretiens la question du (manque de) temps et le sentiment de pression qui pouvait en résulter.
- 53 Défis littéraires, métriques des plateformes, bilans de lecture sur les réseaux sociaux... Toutes ces formes statistiques imposent en effet des plages temporelles bien définies au cours desquelles il faut implicitement lire le plus possible, le plus vite possible — de sorte que la quantification des pratiques littéraires s'exprime fréquemment sous la forme d'un ratio entre une quantité et une unité de temps : nombre de livres lus par an, pages lues par mois, heures consacrées à la lecture par jour, etc. C'est d'ailleurs sûrement en raison de son manque de temps que Gauthier, influenceur littéraire sur Instagram, père de deux jeunes enfants et salarié à temps plein, déclare : « Moi c'est surtout le calcul temporel qui me parle » (40 ans, ingénieur, Instagram). Il utilise une application qui lui permet de « mettre un chronomètre de lecture » pour « calculer le temps » qu'un livre lui prend à lire et d'estimer ainsi le moment où il le finira (voir capture d'écran 4). Pour lui, les pourcentages d'avancée sont inutiles, parce que « si tu

as 600 pages ou tu as 200 pages, 50 % ça reste 50 %. Par contre, c'est pas le même *timing* qu'il y a derrière ».

Capture d'écran 4. Chronomètre de lecture de l'application Book Track, 27 septembre 2025



- 54 De la même façon, plusieurs de nos enquêtés se sont inscrits sur Goodreads en parallèle de leur plateforme littéraire habituelle spécifiquement pour l'une de ses fonctionnalités : le site calcule automatiquement si le rythme de lecture de l'utilisateur est suffisant pour atteindre l'objectif annuel qu'il s'est fixé (voir capture d'écran 5). Par exemple, un lecteur ayant indiqué vouloir lire 52 ouvrages au cours de l'année devra en moyenne enregistrer une lecture par semaine, sous peine d'être déclaré « en retard » par la plateforme. Clara raconte ainsi :

Pour le coup, Goodreads, c'est juste parce qu'il y a ce *challenge* là et qu'il y a vraiment ce truc de « vous êtes en avance » ou « vous êtes en retard ». Donc ça me permet de voir... Parfois il dit « vous avez 2 livres de retard », ou quelque chose comme ça, donc c'est de se dire « ah bah peut être que je vais prendre des plus petits livres pour justement rattraper le retard », entre guillemets. Ou « vous êtes en avance » et ou du coup me dire « bah là je vais prendre un plus gros pavé », par exemple (33 ans, professionnelle de santé, Livraddict).

Capture d'écran 5. Compteur statistique du défi de lecture annuel de la plateforme Goodreads, placé au cœur de son interface



Texte : « 21 livres complétés. 15 livres de retard sur la planification ». Capture d'écran du 27 septembre 2025.

- 55 Une autre pratique fréquente de nos enquêtés consiste à multiplier les supports afin de pouvoir lire à tous les instants de la journée — y compris ceux qui paraissent les moins propices à la lecture. Afin de ne « pas perdre le rythme », Danièle a toujours ainsi trois livres en cours : un ouvrage physique chez elle, une liseuse électronique pour sa pause déjeuner au travail ou ses moments d'attente, et un livre audio pour ses trajets en voiture et le temps consacré aux tâches ménagères (voire, si le livre lui plaît, au moment de sa douche).
- 56 C'est également à l'aune de ce désir d'optimisation du temps de lecture que l'on peut interpréter le refus de certains enquêtés de relire leurs ouvrages. Élodie, qui participe à de nombreux défis littéraires sur Babelio pour lesquels les relectures ne sont pas prises en compte, déclare quant à elle : « Mais non, je relis jamais. Je relis jamais. C'est pas possible. [...] Pas le temps ! Pas le temps. La vie est courte, le nombre de livres que je veux lire est énorme, je peux pas me permettre de relire » (45 ans, conseillère en collectivité territoriale, Babelio). Cette prime à la nouveauté se retrouve par ailleurs sur Instagram, où les sorties récentes paraissent compter davantage que les autres livres, ce qui accentue le sentiment d'urgence ressenti par certains créateurs de contenu : « Il y a ce côté un peu aussi de *fast fashion* du livre où il y a tout le temps des nouveautés [...] et si on découvre un livre plus tard, alors ça a moins de valeur », regrette ainsi Léa, qui gère un compte Instagram littéraire. Elle explique :
- En fait, il y a un côté où il faut montrer [...] qu'on est extrêmement passionné et donc qu'on arrive à trouver du temps tout le temps pour lire. [...] Et puis aussi il y a le fait de lire vite, c'est vu comme quelque chose de positif parce qu'on comprend tout très vite et donc on est moins lent. [...] Mais après ça, c'est peut-être aussi ma vision parce que je culpabilise d'être lente (27 ans, actuellement sans emploi, anciennement libraire, Instagram).
- 57 Si beaucoup de nos enquêtés ont affirmé qu'ils sont vigilants au fait de ne pas (ou plus) laisser les statistiques leur imposer des cadences intenables, persiste tout de même

chez la plupart d'entre eux le désir de maintenir un rythme de lecture soutenu et ininterrompu :

En général, un bouquin, en deux jours c'est plié. [...] Et donc le fait que j'aurais un bouquin de... allez, on va dire 300 pages, qui traînerait sur ma table de salon depuis trois-quatre jours, là je me dirais : « Ouais, peut-être que je vais faire une pause dans celui-là, en lire un autre, parce qu'apparemment t'avances pas » (Danièle).

- 58 Cette pression à lire assidûment trouve peut-être son expression la plus emblématique dans la lutte que mènent de nombreux influenceurs littéraires contre les « pannes de lecture » (ou, selon l'expression anglophone consacrée, « *reading slumps* »). Il serait possible de produire de longs développements sur la conception productiviste et psychologisante de la lecture que ces vidéos promeuvent, mais nous nous contenterons de souligner que ces pannes de lecture sont considérées sur le mode de la maladie à soigner, et qu'elles sont souvent attribuées à la pression que charrient les réseaux sociaux. Comme le dit une Youtubeuse dans une vidéo de conseils pour sortir d'une période de panne de lecture :

Je vous en supplie, rendez-vous service et arrêtez de vous illusionner en pensant que vous pouvez lire 200 livres cette année. [...] Sur les réseaux sociaux, vous allez voir ces bilans impressionnants publiés par le 1 % de la population qui lit réellement autant de livres chaque mois. Mais dans la vraie vie, la plupart des gens qui considèrent la lecture comme une véritable passion ne lisent même pas une fraction de ce nombre, parce que ce n'est tout simplement pas viable. [...] Alors, si vous commencez à vous sentir mal parce que vous ne lisez pas 20 livres par mois [...], vous vous imposez une attente totalement irréaliste. [...] C'est contre-intuitif, mais lorsque vous vous libérez de l'idée que votre panne de lecture vous empêche d'atteindre un objectif de lecture absurde, il sera bien plus facile d'en sortir⁵³.

- 59 Si les discours des lecteurs interrogés se sont rarement révélés aussi extrêmes, certains ont mentionné le sentiment d'anxiété que la quantification faisait peser sur leurs pratiques de la lecture et ses conséquences sur leur santé mentale :

Léa : L'année passée, je voulais lire 365 livres, donc un par jour. Mais en sachant que je lis aussi les mangas, les albums jeunesse tout petits, donc je me disais « je pourrais le faire ». [...] Mais là j'ai envie [...] que ce soit plus positif, en tout cas que l'année passée, donc plutôt entre 100 et 200.

Enquêtrice : Comment ça, « plus positif » ?

Léa : Bah... Avoir moins de pression que je m'impose à moi-même. Enfin, j'ai toujours cette réflexion par rapport au fait que je suis quelqu'un de stressé et que les réseaux sociaux c'est quand même assez... Il y a un côté assez négatif et de pression. Et justement comme on parle de culpabilité, il faut que j'arrive à me détacher de ça et que j'essaye en tout cas de trouver un compromis entre ce que j'ai envie de faire et le bien-être ».

- 60 Ce constat rejoint d'ailleurs l'expérience de certains de nos enquêtés, dont Agathe, qui rapporte avoir été victime de ce qu'elle appelle un « *burn out* littéraire » :

J'[ai] été [sensible aux statistiques] à une époque [...]. À un moment, j'ai tellement eu une *overdose* de *challenges*, de statistiques... enfin de tout ça, que ça m'avait carrément donné une panne de lecture. Et c'est là que je me suis remise en question [...]. Je me suis dit : « Il faut arrêter », parce que vraiment ça me créait un blocage. Parce qu'il y avait trop en fait, c'était limite un *burn out* littéraire.

- 61 Si tous nos enquêtés n'entretiennent pas le même rapport à la mise en nombre de leurs pratiques, il est donc important de noter que les cultures de la quantification — qui invitent à « se *challenge* pour pouvoir maintenir son rythme de lecture » (Clara) —

contribuent aussi parfois à nourrir chez certains lecteurs un rapport anxieux et productiviste aux pratiques littéraires.

Managérialisation du soi lecteur et compétition « contre soi-même »

- 62 Ces cultures de la quantification ont donc pour conséquence indéniable de favoriser une inflation et une intensification, parfois oppressantes, des pratiques littéraires. Il nous faut cependant souligner que, bien loin de l'image stéréotypée de lecteurs hystérisés par la pression des chiffres, la plupart de nos enquêtés ont eu à cœur de nuancer leur rapport aux métriques, prenant par exemple soin de les recontextualiser : « 25 livres de moins qu'en 2022, je suis un poil déçue de ne pas avoir lu davantage mais j'ai été bien occupée, avec une fatigue sur la fin d'année, donc je ne me sens pas totalement coupable » (Agathe).
- 63 De la même façon, en ramenant les défis littéraires à leur statut de jeu, ils en dédramatisent les enjeux : « c'est pas vraiment de la compétition, c'est des adultes qui s'amuse à jouer » (Danièle). Cette mise en perspective des chiffres — rapportés à la contingence de leur contexte de production — se traduit par ailleurs par la condamnation unanime des compétiteurs, dont l'attitude est présentée comme déplacée, voire nuisible. Comme l'expliquait Béatrice, qui organise un *challenge* sur Babelio :
- C'est une compétition avec soi-même. Voilà, c'est à soi-même, je pense, de se fixer des objectifs. Et après, des compétiteurs... J'ai jamais vu quelqu'un dire « Ah ben tiens, tu vois, je t'ai battu, je suis arrivé là avant toi », j'ai jamais vu ça. Et je pense d'ailleurs que j'éliminerais ce *challenger* là si quelqu'un le faisait, tu vois. Je veux pas cet esprit-là dans ce *challenge*.
- 64 À la question de savoir si la mise en nombre de leurs pratiques littéraires induisait chez eux un sentiment de compétition, les personnes interrogées ont de fait très souvent répondu, à l'instar de Miriam : « la seule compétition qu'il y a, c'est vraiment avec moi-même ». Cette auto-référentialité fait écho à ce que Quentin Gilliotte décrit comme le désir « de se distinguer non forcément des autres, mais aussi de son "ancien soi" » (*op. cit.*, p. 143). C'est par la mesure de l'« écart de soi à soi » (Lahire, *op. cit.*, p. 205) — que les métriques prétendent objectiver — que les individus appréhendent donc leurs progrès dans cette entreprise d'optimisation d'eux-mêmes. Cela rejoint le constat de Dominique Cardon, pour qui « les "vérités" statistiques sont devenues instrumentales : ce n'est plus la valeur propre du chiffre qui importe, mais l'évolution de la valeur mesurée entre deux enregistrements » (*op. cit.*, p. 42). Comme le dit Alice :
- Au final, ma seule compétitrice c'est moi-même [...] Enfin, je ne me mets pas de pression spécialement [...]. Mais ouais, le nombre de livres, c'est important. Juste, j'aime bien compter et j'aime bien me comparer à moi-même. [...] J'aime bien voir. « Ok, on est le 1^{er} juillet, j'ai lu 25 livres, je vais voir l'année dernière au 1^{er} juillet ce que j'étais en train de lire. Euh, combien de livres j'avais lu à ce moment-là ».
- 65 Ce rapport aux chiffres fait d'ailleurs écho à ce que Foucault écrit au sujet de l'« examen de conscience », habitude essentielle du « souci de soi » antique :
- Le rapport du sujet à lui-même dans cet examen ne s'établit pas tellement sous la forme d'une relation judiciaire où l'accusé se trouve en face du juge ; il a plutôt l'allure d'une action d'inspection où le contrôleur veut apprécier un travail, une mission accomplie (2014 [1984], p. 86).

- 66 Dès lors, et de façon assez surprenante, nos enquêtés n'ont que ponctuellement fait état d'un sentiment de déception ou de frustration dans les situations où ils ne parvenaient pas à atteindre leurs objectifs chiffrés. Ainsi, pour Gaëlle, « il n'y a pas de culpabilité » à ne pas réussir son défi de lecture : « Ça sert juste comme piqure de rappel, en mode "ah mais j'ai délaissé la lecture en ce moment". » La quasi-totalité des lecteurs interrogés continuent d'ailleurs à se fixer des objectifs quantitatifs même s'ils ne les atteignent pas. Ces échecs sont plutôt pour eux l'occasion de « chercher les causes, pour justement mettre en place des stratégies qui permettraient d'éviter d'arriver au même point l'année suivante » (Élisabeth, 55 ans, conseillère pédagogique, Facebook). Toujours selon les termes de Foucault :

La faute n'est pas réactivée par l'examen pour fixer une culpabilité ou simuler un sentiment de remords, mais pour renforcer, à partir du constat rappelé et réfléchi d'un échec, l'équipement rationnel qui assure une conduite sage (*ibid.*, p. 86-87).

- 67 Nous retrouvons donc ici clairement l'entrelacement du « souci de soi » foucaldien et de l'« agentivité néolibérale⁵⁴ » (Gershon, *op. cit.*) favorisant une conception du « soi comme entreprise⁵⁵ » (McNay, *op. cit.*). Pour nos enquêtés, l'objectif n'est donc pas d'atteindre leurs objectifs à tout prix — ce qui serait la marque d'un esprit de compétition inapproprié —, mais d'objectiver leur propre évolution au travers de leurs performances et de se doter, si elles sont jugées insuffisantes, des moyens de « faire mieux » (Miriam).

L'injonction à la lecture comme substrat à la quantification

- 68 L'un des constats qui structurent cette recherche est par ailleurs que nos enquêtés ont finalement peu critiqué ouvertement la mise en nombre de leurs pratiques littéraires par les plateformes — et ce alors même que les cultures de la quantification sont intimement liées à des rationalités néolibérales qui peuvent paraître, au premier regard, incompatibles avec des pratiques littéraires qui occupent encore une « place très particulière » au sein de la société (Gilliotte, *op. cit.*, p. 319), un espace « à part, soustrait aux règles du monde ordinaire » et souvent sacralisé (Barth-Rabot, *op. cit.*, p. 28). Aucune des personnes interrogées n'a, par exemple, exprimé le désir de quitter la plateforme ou les défis auxquels elle était inscrite. Nous postulons dès lors que la quantification des pratiques littéraires trouve sa raison d'être — ou a minima une solide justification à son expansion — dans la croyance massivement partagée que la lecture est universellement bonne et qu'elle doit être diffusée au plus grand nombre — a fortiori dans le contexte contemporain, où elle est présentée comme une activité en déclin, voire en péril, qu'il s'agirait de sauver⁵⁶.
- 69 Nos enquêtés ont de fait unanimement adopté un discours « prosélyte » (Barth-Rabot, *op. cit.*, p. 76) présentant la lecture comme bonne *en soi*, indépendamment du genre littéraire des livres, de leur format, de leurs auteurs, etc. Élisabeth, ancienne professeure des écoles, expliquait par exemple : « J'ai toujours dit à mes élèves : "Lisez, les amis ! Lisez n'importe quoi, mais lisez". Je leur disais : "Lisez la boîte de céréales le matin, mais lisez, quoi". J'ai un respect total pour la lecture, pour tout type de lecture. » Ces discours s'inscrivent dans le cadre de ce que Marine Siguier appelle l'« injonction intransitive » à la lecture, et qui véhicule l'idée que « peu importe de quelle manière, il faut lire » (*op. cit.*, p. 210, italique dans l'original) — et souvent, ajoutons-nous, lire *plus*. Or, comme nous l'avons précédemment souligné, la mise en nombre d'une pratique

favorise son intensification — ce qui contribue à faire de la quantification des pratiques littéraires un dispositif paradoxalement acceptable.

70 D'ailleurs, et parce qu'ils fonctionnent par incréments de un (on ne peut pas valider de demi-lecture, par exemple), les compteurs de lecture incarnent sûrement la forme la plus concrète de cet idéal d'aplanissement des hiérarchies traditionnelles : ils donnent corps à l'idée que tous les livres *se valent*, faisant de *Guerre et paix* l'équivalent effectif de *Hunger Games* ou de *Gaston Lagaffe*.

71 Il est cependant important de noter que, même si les lecteurs interrogés déclarent quasi unanimement que « l'important c'est de lire, peu importe ce qu'on lit » (Paul, 20 ans, étudiant, Instagram), l'affirmation d'une horizontalité entre tous les livres a souvent ressemblé à un vœu pieu tant nos enquêtés étaient également prompts à remobiliser les hiérarchies classiques du champ littéraire. L'exemple le plus probant est sûrement celui des genres littéraires qui, comme les bandes dessinées (BD), étaient assez unanimement considérés comme « particuliers » — notamment parce qu'ils requièrent moins de temps et « d'efforts » que la lecture d'un roman. Très souvent, nos enquêtés ont donc spontanément précisé s'ils incluaient les bandes dessinées dans le décompte de leurs lectures : « Je lis, je pense, entre 2 à 3 livres par semaine. Enfin, quand je dis livres, c'est BD inclus[es] » (Élise). C'est aussi ce qui explique que plusieurs organisateurs de *challenges* de lecture aient choisi d'imposer des quotas : « par exemple, on a droit à huit bandes dessinées ou romans graphiques et cinq pièces de théâtre » dans l'année (Philippe, 65 ans, retraité, anciennement formateur pour adultes, Babelio).

72 On retrouve ici la conception de la lecture comme ascèse. Comme le note Gilliotte, la lecture est, encore aujourd'hui, la pratique culturelle la plus immédiatement associée à l'habitus bourgeois « qui, pour les pratiques “cultivées” valoris[e] l'autocontrainte, l'effort, la confrontation à une certaine difficulté » (*op. cit.*, p. 295), ce que l'une de nos enquêtés, Anne, synthétise ainsi :

Certaines personnes sacralisent tellement la littérature qu'il faut forcément s'asseoir au calme, dans son salon, dans son bureau, voire dans sa chambre, dans le silence le plus total, prendre le temps de savourer chaque mot, relire trois fois chaque page, bien s'imprégner du style. Par respect pour l'auteur, le temps, la souffrance qu'il a mis à écrire son livre. Il y a vraiment une notion de souffrance chez certains lecteurs et lectrices qui moi me stupéfie toujours. Lire, ça doit faire souffrir.

73 Ce rapport sacrificiel à la lecture passe par la dévaluation des livres jugés trop « faciles » à lire (comme les bandes dessinées ou romans *young adult*), des livres trop courts (certains *challenges* proposant plutôt de se baser sur le nombre de pages lues, unité de mesure jugée « plus juste »), mais également par la stigmatisation des *trop* grands lecteurs. Le cas d'Anne est en ce sens éclairant : « Je lis énormément », « 200 livres par an, voire plus. Je peux même arriver à 300 les années fastes », explique-t-elle, avant de relater les critiques que cela lui a valu de recevoir :

Je m'étais déjà fait attaquer pour le nombre de livres que je lisais par an. [...] J'ai eu droit à : « Lire autant, c'est du gâchis ». Parce que pour cette personne, quand on lit autant, on ne prend pas le temps de savourer le livre, on ne laisse pas le livre vous accompagner, on ne respecte pas l'auteur. [...] Il faut prendre le temps de lire, et lire un roman de 500 pages en deux ou trois jours... Non, ça ne se fait pas. [...] Si tu lis un livre pour te détendre, c'est pas bien. On m'a déjà dit : « De toute façon, toi, tu lis n'importe quoi ». Non, je lis de tout, c'est pas la même chose ! [...] J'ai même entendu dans la vraie vie : « Tu fais quoi ? Tu lis le début, le milieu, la fin ? » Non ! [...] Je lis juste vite.

- 74 Le *trop* grand lecteur est donc immédiatement soupçonné de ne pas respecter l'œuvre et son auteur, de lire des livres de mauvaise qualité ou, dans le cas des défis de lecture, de faire preuve d'un esprit de compétition déplacé, voire de tricher ; en somme, d'être un mauvais lecteur. Cela se retrouve dans les discours des organisateurs de *challenges* littéraires qui, au motif de garantir une forme « d'équité » (Élodie) entre participants, sont souvent à l'affût des cas de fraude :

On sent [quand] la personne triche. On a très vite des doutes. C'est-à-dire que c'est déjà arrivé avec une lectrice qui publiait quatre validations par semaine à peu près. Alors on peut être un très bon lecteur, mais quand tu lis quand même quatre bouquins par semaine — et pas des petits bouquins, hein, pas des BD — au bout d'un moment, ça pose question (Élodie, 45 ans, conseillère en collectivité territoriale, Babelio).

- 75 Il existe donc un seuil au-delà duquel les métriques de lectures cessent d'être impressionnantes et attirent la suspicion (voire l'hostilité) des autres utilisateurs. Il est bien sûr impossible de calculer précisément ce seuil, et c'est d'ailleurs peut-être l'un des traits les plus saillants de nos entretiens : peu importe leurs statistiques et leur intensité, les pratiques de lecture de nos enquêtés donnaient souvent l'impression d'être trop et pas assez dans le même temps. Léa, par exemple, déclare vouloir réduire son objectif de lecture annuel, qu'elle trouve trop dur à tenir, mais a pris comme bonne résolution de lire plus ; Élise considère qu'elle lit « trop », mais explique dans le même temps : « je ne peux pas lire plus vite, mais j'aimerais être capable de lire plus vite, de lire plus », etc.

« Ça doit rester un plaisir ! » : professionnalisation et pathologisation d'une activité de loisir

- 76 L'expression « ça doit rester un plaisir » figure indéniablement parmi les formules les plus spontanément convoquées par nos enquêtés. Si elle traduit certainement la « morale du devoir de plaisir » qui, selon les termes de Pierre Bourdieu, « porte à éprouver comme un échec, propre à menacer l'estime de soi, toute impuissance à "s'amuser" » (1979, p. 424), la récurrence de cette expression exprime peut-être aussi la crainte de voir cet espace de loisir se professionnaliser et ainsi se structurer autour d'impératifs de rentabilité : « Il faut que ça reste un plaisir, quoi. Il faut que le plaisir de lire reste là, sinon ça devient un métier » (Christine, 68 ans, retraitée, anciennement directrice de bibliothèque, Instagram).
- 77 Toutefois, et malgré la fréquence de la formule, le vocabulaire employé par la plupart des autres lecteurs traduisait souvent l'immixtion de logiques managériales dans leurs pratiques littéraires — et l'expression de « *burn out* littéraire » employée par Agathe n'est, en ce sens, sûrement pas tout à fait innocente. Élodie, qui organise (ou coorganise) un grand nombre de challenges littéraires sur Babelio, mobilise d'ailleurs abondamment le champ lexical de la gestion de projet pour décrire son activité :
- Le staff, donc les dix personnes, les acolytes qui m'aident à gérer les challenges, il faut les manager aussi, c'est-à-dire que c'est comme si je manageais une équipe. On a des fils WhatsApp pour pouvoir échanger. On a des visios régulièrement. [...] Là le mois de novembre, c'est un gros mois d'activité pour nous.
- 78 Pour les organisateurs de défis nécessitant de « valider » les lectures des participants, la question des vacances est d'ailleurs un enjeu à part entière. Danièle explique ainsi confier à « [s]es collègues » (c'est-à-dire les utilisateurs qui coorganisent les *challenges*

avec elle) le soin d'assurer cette tâche durant les siennes, ajoutant : « c'est à charge de revanche, puisqu'on part tous en vacances à un moment ou à un autre ». Elle confie également : « Contrairement à la majorité des gens, [...] en vacances, je lis beaucoup moins que quand je travaille, alors que d'autres gens attendent les vacances pour avoir du temps pour lire. »

79 Comme l'a souligné Chris Till au sujet des pratiques sportives, la quantification favorise la transformation des activités de loisir en travail (*labour*) en ce qu'elle permet d'extraire de la valeur (ici, par la vente de données et d'espaces publicitaires) de pratiques qui échappaient jusqu'ici aux logiques de productivité : « Ce "travail gratuit" engendre d'importants profits sans être rémunéré, car les entreprises ont réussi à convaincre les utilisateurs qu'il s'agit de loisir et non de travail en brouillant la frontière⁵⁷ » qui les séparait (2014, p. 449). C'est notamment ce que critique Alain lorsque, parlant de son propre rapport aux statistiques de lecture, il ironise : « Ça participe de cet esprit de compétition. Je suis à 80 % de la réalisation de mon objectif. Mon N+1 va être très content de moi » (60 ans, manager de projets dans le secteur public, Babelio).

80 Nombreux sont donc les enquêtés qui déclarent recourir à des compétences acquises sur leur lieu de travail pour organiser leurs pratiques littéraires — notamment lorsqu'ils utilisent Excel ou des logiciels qui, comme Canva, sont fréquemment employés par les services de communication pour produire du contenu sur les réseaux sociaux : « Je regarde aussi les statistiques, même si j'ai pas le nez dedans, mais on va dire que c'est un peu une déformation professionnelle puisque je travaille dans la communication digitale » (Laura, 35 ans, chargée de communication dans une petite entreprise, YouTube et Instagram). Nombreux sont aussi les enquêtés qui réalisent des *plannings* de leurs lectures et de leurs publications sur les réseaux sociaux. Expliquant son système d'organisation, Paul affirme par exemple : « En soi, si je ne devais plus lire aujourd'hui, le blogue peut fonctionner pendant trois mois sans souci. » Nombreux, finalement, sont ceux qui tiennent une liste des critiques qu'il leur faut écrire ou des items qu'il leur reste à compléter pour leurs *challenges* de lecture, ce qui, avec le temps, devient un automatisme. À la question de savoir s'il lui arrive encore de lire des ouvrages ne lui permettant pas de valider les consignes des défis auxquels il participe, Philippe répond : « Oui ! Oui, c'est le cadeau que je me fais », ajoutant :

Au bout d'un moment, je me dis, « Bon ben là, maintenant, je lis vraiment ce que j'ai envie, sans prendre en compte tous les items de tous les *challenges* auxquels je participe ». Et puis des fois, même une fois que je l'ai terminé, je me dis : « Ah ben tiens, il y a quand même un critère que je peux valider avec ça ». Donc c'est en ce sens que c'est un peu prise de tête et un peu obsessionnel. [...] J'y pense en permanence (Philippe, 65 ans, retraité, anciennement formateur pour adultes, Babelio).

81 Cette notion d'obsession que mentionne Philippe est revenue plusieurs fois chez nos enquêtés, qui parlent également de « charge mentale » (Ambre et Carole, 55 ans, DRH, Babelio) pour décrire le niveau d'organisation qu'exige la participation simultanée à plusieurs défis de lecture. Parmi les 31 lecteurs interrogés, 23 ont ainsi spontanément employé un terme connotant l'addiction, la folie ou une maladie pour qualifier leurs pratiques littéraires : « ça doit être une pathologie hein ! » (Danièle), « je dois te paraître mais complètement psycho ! » (Alice), « je suis dingue » (Miriam), « j'ai beaucoup de copains qui me disent que c'est névrotique », « mon conjoint, il dit que je suis une psychopathe » (Carole), « c'est comme une maladie » (Christine), « c'est un

TOC » (Jeanne, 77 ans, retraitée, anciennement juriste, Babelio), « je suis une malade mentale, moi ! » (Élisabeth), « ça relève de la psychanalyse » (Philippe).

- 82 Cependant, et malgré la tendance de nos enquêtés à associer leurs comportements à des troubles individuels, il nous semble important de souligner qu'ils sont en grande partie entretenus par les plateformes, mais aussi par les acteurs du monde de l'édition, qui bénéficient largement du travail de promotion des lecteurs disposant de blogues ou de comptes sur les réseaux sociaux et qu'ils rémunèrent, dans une immense majorité des cas, exclusivement en envois de livres. Cette tendance est particulièrement visible dans le cas des influenceurs Instagram interrogés. Florence déclare par exemple : « ça m'interpelle de voir depuis quelques années cette pile qui grimpe plus vite que ce que je lis alors que je lis plus qu'avant aussi », ajoutant qu'elle cherche à comprendre « pourquoi [elle se] laisse dépasser par cette pile à lire » qu'elle appelle son « Everest à lire ». La réponse est sûrement, dans son cas, à chercher du côté des colis presse : en 3 semaines, elle a reçu 13 livres, dont une partie était envoyée par des éditeurs à qui elle n'avait même pas communiqué son adresse. Pourtant, Florence affirme que sa relation avec les maisons d'édition n'« est pas commerciale. Parce que moi je ne suis pas payée pour ce que je fais et ça prend du temps », précisant que demander d'être payée pour son travail « n'est même pas une chose qui [lui] viendrait à l'idée de demander ». Elle explique : « Moi j'estime que c'est pas commercial et que c'est plus un engagement peut-être moral ou en tout cas, c'est plus sincère et authentique que du commercial ».
- 83 Dans ce cadre, la propension de nos enquêtés à rappeler la prévalence de leur plaisir ainsi que la sincérité et l'authenticité de leur démarche tient peut-être aussi à la conscience qu'ils ont de la difficulté de maintenir des pratiques improductives au sein d'espaces structurés par la quantification.

Conclusion

- 84 Notre objectif était de jeter ici les bases d'une analyse des effets des cultures de la quantification lorsqu'elles s'appliquent à la pratique culturelle « la plus hautement symbolique et la plus investie de valeurs » (Barth-Rabot, *op. cit.*, p. 9). De fait, et contrairement à d'autres formes de *tracking*, la quantification des pratiques littéraires n'implique pas seulement d'optimiser ses métriques, mais aussi de donner, dans le même temps, des gages de son désintéressement — le bon lecteur lit *plus*, et *assidûment*, mais ne lit pas *trop* ; il adopte la posture ascétique nécessaire à la juste appréciation des œuvres et du travail de l'auteur ; il réaffirme sans cesse la prévalence de son plaisir, la gratuité de son geste, le désintéressement de son amour de l'art.
- 85 Ces discours ne sont bien sûr pas neutres et servent une posture socioculturelle : comme le rappelle Barth-Rabot, « la lecture est une pratique sur laquelle les pratiquants parlent en général assez volontiers et facilement », mais « [l]a valeur symbolique de la lecture et la dimension identitaire d'une pratique qui est souvent aussi un positionnement, ont des implications méthodologiques » (*op. cit.*, p. 21). C'est pour cette raison que notre recherche visait également à analyser les pratiques concrètes, quotidiennes des lecteurs immergés au sein des cultures de la quantification : leurs façons de lire et d'interagir avec les interfaces, leurs rituels, leurs habitudes d'achat de livres, l'organisation de leur temps et de leurs espaces, etc. C'est l'écart qui existait parfois entre les discours de nos enquêtés — structurés par une

conception à la fois noble et affective de ce que lire veut dire, un idéal démocratique de la lecture comme loisir nécessairement bénéfique et plaisant — et la réalité matérielle de leurs pratiques — tournées vers une optimisation du soi parfois anxiogène, empruntant à l'entreprise néolibérale son ethos compétitif et sa grammaire managériale — que nous avons ici voulu mettre en lumière.

BIBLIOGRAPHIE

- AKRICH Madeleine (1987), « Comment décrire les objets techniques ? », *Techniques et culture*, 9, p. 49.
- ALBRECHTSLUND Anders et Peter LAURITSEN (2013), « Spaces of everyday surveillance: Unfolding an analytical concept of participation », *Geoforum*, 49, p. 310-316.
- ALONSO William et Paul STARR (dir.) (1987), *The politics of numbers*, New York, Russell Sage Foundation.
- ANNABELL Taylor et Nina Vindum RASMUSSEN (2024), « Spotify (Un)wrapped: How ordinary users critically reflect on Spotify's datafication of the self within creative workshops », *Journal of Gender Studies*, p. 1-17, DOI : <https://doi.org/10.1080/09589236.2024.2433674>.
- BARTH-RABOT Cécile (2023), *La lecture. Valeur et déterminants d'une pratique*, Paris, Armand Colin.
- BEER David (2015), « Productive measures: Culture and measurement in the context of everyday neoliberalism », *Big Data et Society*, 2(1), <https://doi.org/10.1177/2053951715578951>.
- BEER David (2016), *Metric Power*, Londres, Palgrave Macmillan.
- BERMAN Elizabeth et Daniel HIRSCHMAN (2018), « The sociology of quantification: Where are we now? », *Contemporary Sociology: A Journal of Reviews*, 47(3), p. 257-266.
- BESSARD-BANQUY Olivier (2022), *Splendeurs et misères de la littérature. Ou la démocratisation des lettres, de Balzac à Houellebecq*, Paris, Armand Colin.
- BÖHM Steffen et Chris LAND (2009), « No measure for culture? Value in the new economy », *Capital et Class*, 33(1), p. 75-98.
- BOURDIEU Pierre (2015), *Les règles de l'art. Genèse et structure du champ littéraire*, Paris, Points.
- BOURDIEU Pierre (1979), *La Distinction*, Minuit.
- BROWN Wendy (2017), *Undoing the demos: Neoliberalism's Stealth Revolution*, Princeton, Zone Books.
- CALVIGNAC Cédric (2021), « Traductions sociotechniques des principes axiologiques du *quantified self*. Analyse d'un corpus de brevets US dédiés à la mesure et à la gestion du sommeil », *Réseaux*, 228(4), p. 131-169.
- CANDEL Etienne, Valérie JEANNE-PERRIER et Emmanuël SOUCHIER (2012), « Petites formes, grands desseins. D'une grammaire des énoncés éditoriaux à la standardisation des écritures », dans Jean DAVALLON, *L'Économie des écritures sur le web*, p. 165-201.
- CARDON Dominique (2015), *À quoi rêvent les algorithmes. Nos vies à l'heure des big data*, Paris, Seuil.

- CHARTIER Roger (2015), *La main de l'auteur et l'esprit de l'imprimeur. XVI^e-XVIII^e siècle*, Paris, Gallimard.
- CITTON Yves (2016), « Subjectivations computationnelles à l'ère numérique », *Multitudes*, 62(1), p. 45-64.
- CITTON Yves (2017), *Médiarchie*, Paris, Seuil.
- CROSBY Alfred (2003), *La mesure de la réalité. La quantification dans la société occidentale, 1250-1600*, Paris, Éditions Allia.
- DAGIRAL Éric, Séverine DESSAJAN, Tomas LEGON, Olivier MARTIN, Anne-Sylvie PHARABOD et Serge PROULX (2019), « Faire place aux chiffres dans l'attention à soi. Une sociologie des pratiques de quantification et d'enregistrement aux différents âges de la vie », *Réseaux*, 216(4), p. 119-156.
- DAGIRAL Éric, Christian LICOPPE, Olivier MARTIN et Anne-Sylvie PHARABOD (2019), « Le Quantified Self en question(s). Un état des lieux des travaux de sciences sociales consacrés à l'automesure des individus », *Réseaux*, 216(4), p. 17-54.
- DALGALARRONDO Sébastien et Tristan FOURNIER (2019), « Introduction. Les morales de l'optimisation ou les routes du soi », *Ethnologie française*, 49(4), p. 248.
- DARDOT Pierre et Christian LAVAL (2010), *La nouvelle raison du monde. Essai sur la société néolibérale*, Paris, La Découverte.
- DAY Sophie, Celia LURY et Nina WAKEFORD (2014), « Number ecologies: Numbers and numbering practices », *Distinktion: Journal of Social Theory*, 15(2), p. 123-154.
- DE CRISTOFARO Diletta et Simona CHIODO (2023), « Quantified sleep: Self-tracking technologies and the reshaping of 21st-Century subjectivity », *Historische Sozialforschung = Historical Social Research*, 48(2), p. 176-193.
- DESEILLIGNY Oriane (2014), « La recommandation sur le Web : entre héritages formels et logiques comptables », *Communication et langages*, 179(1), p. 33-47.
- DESROSÈRES Alain (2008), *Pour une sociologie historique de la quantification. L'argument statistique I*, Paris, Presses des Mines.
- DESROSÈRES Alain (2010), *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte.
- DONNAT Olivier (2020), « À propos de la quantification de la vie culturelle » [billet], *Politiques de la culture*, 21 octobre, <https://chmcc.hypotheses.org/11138>, page consultée le 27 novembre 2025.
- DRUCKER Johanna (2020), *Visualisation : l'interprétation modélisante*, traduction de Marie-Mathilde BORTOLOTTI, Montreuil, Éditions B42.
- DUBOIS Vincent (2012), *La politique culturelle : genèse d'une catégorie d'intervention publique*, Paris, Belin.
- DUBOIS Vincent (2015), « The politics of cultural statistics: From legitimisation to critique (France 1960-1990) », dans Lachlan MACDOWALL, Marnie BADHAM, Emma BLOMKAMP et Kim DUNPHY (dir.), *Making culture count. The politics of cultural measurement*, Londres, Palgrave Macmillan UK, p. 42-52.
- ESPELAND Wendy Nelson et Mitchell L. STEVENS (2008), « A Sociology of Quantification », *European Journal of Sociology/Archives européennes de sociologie*, 49(3), p. 401-436.

- ESPELAND Wendy et Michael SAUDER (2007), « Rankings and reactivity: How public measures recreate social worlds », *American Journal of Sociology*, 113(1), p. 1-40.
- ESPELAND Wendy et Mitchell L. STEVENS (1998), « Commensuration as a social process », *Annual Review of Sociology*, 24(1), p. 313-343.
- FOUCAULT Michel (2001), *Dits et écrits 2 (1976-1988)*, Paris, Gallimard.
- FOUCAULT Michel (2004 [1979]), *Naissance de la biopolitique. Cours au Collège de France (1978-1979)*, Paris, Gallimard.
- FOUCAULT Michel (2014), *Histoire de la sexualité. Tome 3 : Le souci de soi*, Paris, Gallimard.
- GANE Nicholas (2014), « The emergence of neoliberalism: thinking through and beyond Michel Foucault's lectures on biopolitics », *Theory, Culture & Society*, 31(4), p. 3-27.
- GERSHON Ilana (2011), « Neoliberal agency », *Current Anthropology*, 52(4), p. 537-555.
- GILLIOTTE Quentin (2022), *L'expérience culturelle en régime numérique : explorer, ranger et consommer*, Paris, Mines Paris Tech/PSL.
- GROSSER Ben (2014), « What do metrics want? How quantification prescribes social interaction on Facebook », *Computational Culture*, 4, <http://computationalculture.net/what-do-metrics-want/>, page consultée le 10 octobre 2025.
- HACKING Ian (1990), *The taming of chance*. 17, Cambridge, Cambridge University Press.
- HACKING Ian (2006), « Making Up People », *London Review of Books*, 28(16).
- JEANNERET Yves et Emmanuël SOUCHIER (2005), « L'énonciation éditoriale dans les écrits d'écran », *Communication et langages*, 145(1), p. 3-15.
- LAHIRE Bernard (2006), *La culture des individus. Dissonances culturelles et distinction de soi*, Paris, La Découverte.
- LAVAL Christian (2014), « Chapitre 10. L'entreprise comme nouvelle forme de gouvernement. Usages et mésusages de Michel Foucault », dans *Usages de Foucault*, Paris, Presses universitaires de France, p. 143-158.
- LOPEZ-GONZALEZ Hibi (2014), « Quantifying the immeasurable: A reflection on sport, time and media », *Journal of the Philosophy of Sport*, 41(3), p. 347-362.
- LUPTON Deborah (2015), « Quantified sex. A critical analysis of sexual and reproductive self-tracking using apps », *Culture, Health et Sexuality*, 17(4), p. 440-453.
- LUPTON Deborah (2016), *The quantified self. A sociology of self-tracking*, Cambridge, Polity Press.
- LUPTON Deborah (2020), *Data selves: More-than-human perspectives*, Cambridge, Polity Press.
- MACDOWALL Lachlan, Marnie BADHAM, Emma BLOMKAMP et Kim DUNPHY (dir.) (2015), *Making culture count. The politics of cultural measurement*, Londres, Palgrave Macmillan.
- MARTIN Olivier (2020), *L'empire des chiffres*, Paris, Armand Colin.
- MAUGER Gérard et Claude POLIAK (1998), « Les usages sociaux de la lecture », *Actes de la recherche en sciences sociales*, 123(1), p. 3-24.
- MCNAY Lois (2009), « Self as enterprise: Dilemmas of control and resistance in Foucault's *The Birth of Biopolitics* », *Theory, Culture et Society*, 26(6), p. 55-77.
- MENNICKEN Andrea et Robert SALAIS (dir.) (2022), *The new politics of numbers: Utopia, evidence and democracy*, New York, Springer International Publishing.

- MULLER Jerry Z. (2018), *The tyranny of metrics*, Princeton, Princeton University Press.
- MURRAY Simone (2018), *The digital literary sphere: Reading, writing, and selling books in the Internet Era*, Baltimore, Johns Hopkins University Press.
- MURRAY Simone (2021), « Secrets agents: Algorithmic culture, Goodreads and datafication of the contemporary book world », *European Journal of Cultural Studies*, 24(4), p. 970-989.
- PECK Jamie (2010), *Constructions of neoliberal reason*, Londres, Oxford University Press.
- PHARABOD Anne-Sylvie (2019), « “Faire ses 10 000 pas”, vraiment ? Une enquête sur les pratiques de self-tracking ordinaires », *Réseaux*, 216(4), p. 157-187.
- PHARABOD Anne-Sylvie, Véra NIKOLSKI et Fabien GRANJON (2013), « La mise en chiffres de soi. Une approche compréhensive des mesures personnelles », *Réseaux*, 177(1), p. 97-129.
- PORTER Theodore (2017), *La confiance dans les chiffres. La recherche de l'objectivité dans la science et dans la vie publique*, traduction de Gérard MARINO, Paris, Les Belles Lettres.
- PRITZ Sarah Miriam (2016), « Making emotions count: The self-tracking of feelings », dans Stefan SELKE (dir.), *Lifelogging*, New York, Springer, p. 179-189.
- RICCI Francesco, Li'or RO'KAH et Bracha SHAPIRA (2022), *Recommender systems handbook*, New York, Springer.
- ROSE Nikolas (1988), « Calculable minds and manageable individuals », *History of the Human Sciences*, 1(2), p. 179-200.
- ROSE Nikolas (1991), « Governing by numbers: Figuring out democracy », *Accounting, Organizations and Society*, 16(7), p. 673-692.
- RUCKENSTEIN Minna (2014), « Visualized and interacted life: Personal analytics and engagements with data doubles », *Societies*, 4(1), p. 68-84.
- RUCKENSTEIN Minna et Natasha Dow SCHÜLL (2017), « The datafication of health », *Annual Review of Anthropology*, 46(1), p. 261-278.
- SELKE Stefan (dir.) (2016), *Lifelogging*, New York, Springer.
- SEYFERT Robert et Jonathan ROBERGE (2016), « What are algorithmic cultures? », dans Robert SEYFERT et Jonathan ROBERGE, *Algorithmic cultures: Essays on meaning, performance and new technologies*, Londres, Taylor et Francis, p. 1-25.
- SHORE Cris et Susan WRIGHT (2024), *Audit culture*, Londres, Pluto Press.
- SIGUIER Marine (2025), *Des livres et des likes. Représentations de la lecture sur les réseaux sociaux*, Paris, Sorbonne Université Presses.
- STENGER Thomas et Alexandre COUTANT (2009), « La prescription ordinaire sur les réseaux socionumériques », *Médias 09, entre communautés et mobilité*, 14 et 15 décembre, Paris, p. 1-24.
- STRIPHAS Ted (2009), *The Late Age of print. Everyday book culture from consumerism to control*, Colombia, Columbia University Press.
- STRIPHAS Ted (2010), « The abuses of literacy: Amazon Kindle and the right to read », *Communication and Critical/Cultural Studies*, 7, p. 297-317.
- SUPIOT Alain (2015), *La gouvernance par les nombres. Cours au Collège de France (2012-2014)*, Paris, Fayard.
- THOMÉ Cécile (2023), « Quantifier le corps menstrué. Étude des usages des applications de suivi du cycle », *Réseaux*, 241(5), p. 275-314.

TILL Chris (2014), « Exercise as labour: Quantified self and the transformation of exercise into labour », *Societies*, 4(3), p. 446-462.

VOLPILHAC Aude (2015), « *Le secret de bien lire* ». *Lecture et herméneutique de soi en France au XVII^e siècle*, Paris, Honoré Champion.

WHITSON Jennifer R. (2013), « Gaming the quantified self », *Surveillance et society*, 11(1/2), p. 163-176.

NOTES

1. Traduction de l'autrice. Citation originale : « hallowed respite from cynically surveilling social media ».
2. D'une durée d'une heure à trois heures.
3. Lise, 38 ans, assistante dans une association, Instagram. Tous les enquêtés ont été pseudonymisés.
4. Alain, 60 ans, manager de projets dans le secteur public, Babelio.
5. Traduction de l'autrice. Citation originale : « still very far from having general claims or a common theoretical language ».
6. Traduction de l'autrice. Citation originale : « the production and communication of numbers ».
7. Traduction de l'autrice. Citation originale : « social study of number [...] of what numbering does, rather than what numbering is ».
8. Traduction de l'autrice. Citation originale : « everyday neoliberalism ».
9. Traduction de l'autrice. Citation originale : « systems of measurement and competition become a routine part of everyday attitudes, imaginaries and practices ».
10. Traduction de l'autrice. Citation originale : « number ecologies ».
11. Traduction de l'autrice. Citation originale : « living in and with numbers ».
12. Traduction de l'autrice. Citation originale : « Self-tracking practices are directed at regularly monitoring and recording, and often measuring, elements of an individual's behaviours or bodily functions. »
13. Traduction de l'autrice. Citation originale : « making up people ».
14. Traduction de l'autrice. Citation originale : « algorithmic cultures ».
15. Traduction de l'autrice. Citation originale : « I use the term 'cultures' to encapsulate the view [...] that the practices, meanings, discourses and technologies associated with self-tracking are inherently and inevitably the product of broader social, cultural and political processes ».
16. Traduction de l'autrice. Citation originale : « everyday cultural forms ».
17. La définition du défi que nous donnons ici provient du blog de lecture de Light and Smell : <https://lightandsmell.wordpress.com/2025/09/12/pal-week-end-a-1000-session-personnelle-2025/>. Consulté en novembre 2025.
18. Kendra Albert, « The Great Book Project of 2013 », *Quantified Self*, <https://quantifiedself.com/show-and-tell/?project=521>, page consultée le 14 avril 2025 ; Kyrill Potapov, « Finding My Optimum Reading Speed », *Quantified Self*, <https://quantifiedself.com/show-and-tell/?project=989>, page consultée le 14 avril 2025.
19. Traduction de l'autrice. Citation originale : « Neoliberalism doesn't exist in any pure form ».
20. Traduction de l'autrice. Citation originale : « too loose a signifier ».
21. Traduction de l'autrice. Citation originale : « The concept of neoliberalism might provide a starting point, but it certainly should not be a conclusion in its own right. »
22. Traduction de l'autrice. Citation originale : « Neoliberal agency ».
23. Traduction de l'autrice. Citation originale : « Self as enterprise ».

24. Traduction de l'autrice. Citation originale : « reflexively manages oneself as though the self was a business ».
25. Si nous employons ici le masculin neutre, il nous semble cependant important de rappeler que la lecture est une activité majoritairement féminine, y compris dans le panel de personnes ici interrogées.
26. Recrutées par les messageries privées de leurs plateformes ou par courriel.
27. Neuf d'entre elles avaient également un blogue et cinq utilisaient Goodreads (première plateforme de recommandation littéraire anglophone).
28. En France, le Centre national du livre définit comme « grands lecteurs » toutes les personnes lisant plus de vingt livres par an — ce qui représente 31 % des Français.
29. Une partie de nos enquêtés a ainsi été recrutée par l'intermédiaire du projet *Lisez-vous le Belge ?*, destiné à promouvoir la littérature de Belgique francophone.
30. Bookstagram est le nom communément donné aux communautés littéraires sur Instagram.
31. Ces fichiers sont cependant en « lecture seule » : ils ne sont modifiables que par les administrateurs du défi, qui disposent dès lors du pouvoir d'inscrire au registre les lectures « validées »... ou de les refuser.
32. Traduction de l'autrice. Citation originale : « Recommender systems are software tools and techniques providing suggestions for items to be of use to a user. »
33. Comme le site *My year in books*, créé par Éric Li : <https://www.myyearinbooks.com/>, page consultée le 14 avril 2025.
34. Voir par exemple le quiz de la campagne *L'heure de lire*, lancée par le média belge *Le Vif*, qui conclut : « En réduisant votre temps d'écran de 10 %, vous pouvez lire 24 minutes par jour. Cela représente 101,25 pages par semaine, et environ 16,8 livres par an », <https://www.levif.be/lheuredelire/>, page consultée le 14 avril 2025.
35. Traduction de l'autrice. Citation originale : « Numbers are easy to dislodge from local contexts and reinsert in more remote contexts. »
36. Traduction de l'autrice. Citation originale : « Where a number comes from and where it goes matters for what that number is and does. »
37. Traduction de l'autrice. Citation originale : « For example, would we add as many friends if we weren't constantly confronted with how many we have? »
38. Traduction. Citation originale : « Once users engage in these platforms, both voluntary and involuntary self-tracking operates. »
39. Sur Babelio et Livraddict, il est possible d'indiquer un livre comme « abandonné », mais il ne compte pas dans les statistiques de lecture annuelles.
40. Destiny Sidwell (2024), *Reviewing the 148 books I read in 2023 in a few sentences — 2023 wrap up*, vidéo en ligne, <https://www.youtube.com/watch?v=bac2P1vbzsY>, page consultée le 14 avril 2025. Traduction de l'autrice. Citation originale : « My reading, my rules. If I read more than 50 %, definitely 60 % of a book, and then end up “DNF-ing” [NDLA : DNF est l'acronyme de « did not finish »], I still count it as a book read because I spent a lot of time reading the book, and since it's my personal reading I count it. »
41. Règle du *challenge* littéraire « Programme Progress and PAL », sur Livraddict, « [Illimité] Programme Progress and PAL 2.0 », <https://www.livraddict.com/forum/viewtopic.php?p=1&id=37992>, page consultée le 14 avril 2025.
42. Notons tout de même que Babelio a annoncé en mars 2025 la possibilité d'indiquer une relecture sur le site, après avoir expliqué, sur son compte Threads, que cette fonctionnalité avait été demandée « en majorité » par des lecteurs qui « voulaient que ça compte dans leur défi de l'année de lecture en cours ». Publication du 8 janvier 2024, disponible à l'adresse : https://www.threads.com/@babelio_/post/C11roBtsoFc?hl=fr.
43. Traduction de l'autrice. Citation originale : « it is useful to think [...] metrics as having built-in intentions ».

44. Traduction de l'autrice. Citation originale : « I would argue that metrics are no less immune to the desire for more than we are — that, as an instrument of capital, the metric is just as susceptible to capital's growth fetish. Metrics want to “pile up”, to be larger, to grow higher. »
45. Voir par exemple la vidéo : Ellen Catherine (2023), *HUGE unboxing book haul. 60 + books*, <https://www.bing.com/videos/search?q=HUGE+unboxing+book+haul.+60+%2b+books&qpv=HUGE+unboxing+book+haul.+60+%2b+books&FORM=VDRE>, page consultée le 14 avril 2025.
46. Voir Destiny Sidwell (2024), *Reviewing the 148 books I read in 2023 in a few sentences —2023 wrap up*, vidéo en ligne, page consultée le 14 avril 2025.
47. Paolo Z (2020), *J'ai lu 30 livres en 30 jours pour guérir un complexe d'infériorité*, vidéo en ligne, page consultée le 14 avril 2025.
48. LaceyBooklovers (2022), *How I Read Almost 500 Books In A Year | My Tips & Tricks To Read More Books !*.
49. Gennys (2024), *600 € de livres ? BIG BOOK HAUL 2024 (Romance, Dark Romance, Fantasy, Thriller)*, <https://www.youtube.com/watch?v=2RXsdKz371s>, page consultée le 14 avril 2025.
50. Lilo (2024), *50 LIVRES : BOOK HAUL*, <https://www.youtube.com/watch?v=LE0ICu3NiVo&pp=ygUJYm9vayBoYXVs>, page consultée le 14 avril 2025.
51. Nina Quill (2024), *AU SECOURS / BOOK HAUL + 100 LIVRES*, <https://www.youtube.com/watch?v=WJlTFHkq36U>, page consultée le 14 avril 2025.
52. Cette expression est utilisée pour qualifier les livres qu'une personne possède, mais n'a pas encore lus.
53. Traduction de l'autrice. lexi aka newlynova (2023), *these books will break your reading slump*, https://www.youtube.com/watch?v=_BySqbalBEE, page consultée le 14 avril 2025.
54. Traduction de l'autrice. Citation originale : « Neoliberal agency ».
55. Traduction de l'autrice. Citation originale : « Self as enterprise ».
56. Notons cependant que ces discours n'ont rien de nouveau (Bessard-Banquy, 2022).
57. Traduction de l'autrice. Citation originale : « This “free labour” generates vast amounts of income, but is not paid, because the corporations have successfully convinced users that it is leisure, not labour, through an erosion of the distinction between work and play. »

RÉSUMÉS

Si l'omniprésence des chiffres au sein des sociétés occidentales a été largement documentée, rares sont les travaux qui se sont spécifiquement intéressés aux effets de la mise en nombre des pratiques culturelles sur les individus. Afin de caractériser les spécificités de la quantification lorsqu'elle s'applique à une activité aussi socialement valorisée — voire sacralisée — que les pratiques littéraires, nous avons conduit 31 entretiens semi-dirigés avec des lecteurs produisant ou consultant régulièrement du contenu littéraire sur Babelio, Livraddict ou Instagram — trois plateformes numériques ayant pour trait commun de faire figurer, au cœur de leurs interfaces, une multitude d'artefacts quantifiés : compteurs, statistiques, classements, pourcentages, graphiques, notes, etc. Notre objectif est ainsi de souligner la façon dont ce que nous appelons les « cultures de la quantification » favorisent une inflation des pratiques littéraires et un rapport managérial au soi lecteur. Cependant, et contrairement à d'autres formes de *tracking*, la quantification des pratiques littéraires n'implique pas seulement de maximiser ses métriques : le

bon lecteur lit *plus*, et *assidûment*, mais ne lit pas *trop* — ce qui serait perçu comme un rapport déplacé, compétitif ou consumériste à la littérature. En somme, nous souhaitons montrer comment la quantification des pratiques littéraires favorise des discours et des pratiques souvent contradictoires, poussant les lecteurs à faire croître leurs métriques tout en manifestant une indifférence à leur égard, les invitant à augmenter la cadence tout en adoptant une posture ascétique, à réaffirmer la prévalence du plaisir de lecture tout en empruntant à l'entreprise néolibérale son ethos compétitif et sa grammaire managériale.

Despite the extensive body of work on the pervasiveness of quantification in Western societies, very few studies have examined the effects of translating cultural practices into numbers on individuals. In order to identify the specific features of quantification when applied to an activity that is as socially valued—and even sacralized—as literary practices, we conducted 31 semi-structured interviews with readers who regularly produce or consult literary content on Babelio, Livraddict, or Instagram—three digital platforms that share a common characteristic: their interfaces prominently display a multitude of quantified artifacts, such counters, statistics, rankings, percentages, graphs, ratings, etc. Our goal is to highlight how what we term “cultures of quantification” foster an inflation of literary practices and a managerial relationship with the reader. However, unlike other forms of tracking, the quantification of literary practices does not simply mean maximizing metrics: a good reader reads *more*, and *regularly*, but not *too much*—otherwise it would be seen as an inappropriate, competitive, or consumerist relationship with literature. In short, we aim to show how the quantification of literary practices fosters often contradictory discourses and behaviours, pushing readers to increase their metrics while expressing indifference toward them, inviting them to read faster while adopting an ascetic stance, and to reaffirm the primacy of reading pleasure while borrowing the competitive ethos and managerial grammar of neoliberal enterprise.

A pesar de la magnitud del corpus de los trabajos que conciernen la omnipresencia de la cuantificación, en el marco de las sociedades occidentales, pocas investigaciones se han interesado hasta ahora en los efectos del número de prácticas culturales de las personas. Con el fin de describir las especificidades de la cuantificación, cuando se aplica a una actividad también socialmente valorada — incluso sacralizada — como las prácticas en los escritos, hemos llevado a cabo 31 entrevistas semidirigidas con lectores productivos o que consultan regularmente contenidos en Babelio, Livraddict o Instagram — tres plataformas digitales que tienen como característica común hacer aparecer, en el marco de sus interfaces, una multitud de artefactos cuantificados: contadores, estadísticas, clasificaciones, porcentajes, gráficos, notas, etc. Nuestro objetivo es resaltar de esta manera lo que denominamos “culturas de la cuantificación” que favorecen un aumento de las prácticas en los escritos y un informe de gestión al lector. Sin embargo, a diferencia de otras formas de *tracking*, la cuantificación de las prácticas en los escritos no implica solamente optimizar sus indicadores: el buen lector lee *más* y asiduamente, pero no lee demasiado, lo que sería percibido como una relación inapropiada, competitiva o consumista de los escritos. En pocas palabras, aspiramos demostrar de qué manera la cuantificación de las prácticas en los escritos favorece los discursos y prácticas a menudo contradictorias, llevando a los lectores a incrementar sus indicadores, al mismo tiempo que manifiestan indiferencia en su contra y los invita a aumentar la frecuencia, al mismo tiempo que se adopta una posición ascética, a reafirmar la prevalencia del placer en la lectura, al mismo tiempo que se toma de la empresa neoliberal su ethos competitivo y su gramática de gestión.

INDEX

Palabras claves : data, lectura, plataformas, usuarios, Foucault

Keywords : data, reading, platforms, users, Foucault

Mots-clés : plateformes, lecture, utilisateurs, data, Foucault

AUTEUR

JULIETTE PARMENTIER

Juliette Parmentier est doctorante en sciences de l'information et de la communication à l'Université libre de Bruxelles (ReSIC) et en interaction humain-machine au sein de la faculté d'informatique de l'Université de Namur (NADI). Courriel : juliette.parmentier@ulb.be

Lutter contre les violences des données issues de la datafication

Pour une littératie des données queer

Ugo Verdi

- 1 La *datafication*, en tant que « quantification des interactions sociales et de leur transformation en données numériques » (Carlson *et al.*, 2011) et comme « transformation de l'action sociale en données quantifiées en ligne permettant le traçage en temps réel et l'analyse prédictive » (Mayer-Schönberger et Cukier, 2013), est un phénomène dont les conséquences négatives sont décriées au sein des études critiques sur les données (*critical data studies*). Elles enquêtent sur les données selon un angle d'analyse devant permettre le décryptage des collectes, des analyses et des utilisations (Richterich, 2018), s'attelant à déconstruire les mythes du néo-positivisme numérique (Mosco, 2014), en particulier les promesses liées aux analyses quantitatives issues du *big data*. Et au-delà, de remettre en question le pouvoir — renvoyant ici à « la multiplicité des rapports de force immanents au domaine où ils s'exercent, et sont constitutifs de leur organisation » (Foucault, 1994, p. 121-122) — sur les données et son inégale répartition.
- 2 L'essentialisation des données en « une » donnée unique et uniforme amène à une généralisation des pratiques de collecte et de traitement (Guyan, 2022). Or, les méthodologies utilisées ont des conséquences politiques, nécessitant ici un regard politique et épistémique (Proulx, 2020) du fait que les données sont chargées d'un « enjeu politique fort » (Théviot, 2023) et détentrices d'un certain pouvoir (D'Ignazio, 2017).
- 3 Dans ce cadre, les personnes *queer* subissent une triple violence. Tout d'abord, une invisibilisation de leurs problématiques spécifiques, ce que D'Ignazio (2024) nomme les « données manquantes » (*missing data*) du fait que lesdites données sont négligées et ne font pas l'objet de priorisation, de collecte, de conservation et de publication par les institutions, et ce malgré les demandes politiques selon lesquelles elles devraient être collectées et rendues disponibles, comme dans le cas des féminicides étudiés par le « féminisme des données ». C'est ce qui a été soulevé par Ijjasz-Vasquez et Cortez en

2017 qui parlaient d'un « fossé des données LGBTI » (*data gap LGBTI*) à combler. Puis, une essentialisation de leurs caractéristiques où l'entraînement des algorithmes et des intelligences artificielles (IA) sur des jeux de données généralistes a produit des résultats biaisés. Enfin, des stigmatisations résultant de mauvais calibrage des jeux de données et des algorithmes associés, ainsi que de mauvaises utilisations des résultats issus de ces derniers : si les discriminations issues des données et des algorithmes ont été souvent répertoriées dans la littérature (Belcastro *et al.*, 1994 ; Cervulle et Freitas, 2015 ; D'Ignazio et Klein, 2020 ; Noble, 2018 ; O'Neill, 2017), c'est moins le cas vis-à-vis des données *queer*. Aussi, comment proposer des formes alternatives de données pour représenter au mieux les vies *queer* ayant été discriminées du fait de la *datafication* où règne un certain « colonialisme des données¹ » (Couldry et Mejias, 2019) ?

Objectif de l'article

- 4 L'objectif de cet article est de poser les bases d'une littératie des données *queer* ayant pour but de rendre visible l'ensemble des philosophies et des méthodes critiques et *queer* pour permettre la création, l'analyse et la représentation de données *queer* plus représentatives des vies *queer*. Il prend appui sur une synthèse d'approches employées dans la conception et l'analyse des jeux de données issues des littératies *queer* et critiques (des données). Il sert de liant à des courants de recherches souvent silotés et promeut la synergie entre les acteurs, toutes disciplines confondues, qui travaillent sur des approches et des thématiques spécifiques (ex. : le corps, l'espace, le temps, les discours, etc.) afin de réfléchir collectivement sur ces dernières et défricher de nouveaux espaces de recherches.
- 5 La littératie des données *queer* se veut être une littératie des données (*data literacy*) dite « spécialisée » (Verdi, 2023), à savoir une littératie reconfigurée pour répondre à une problématique précise, ce que ne permet pas une vision classique de ladite littératie des données. La littératie des données est la capacité « éthique et critique d'accéder, d'interpréter, d'évaluer, de gérer, de manipuler, de créer et d'utiliser les données grâce à une connaissance de leurs caractéristiques, de leur cycle de vie et des effets engendrés par leur usage » (Verdi, 2023). Dans le cadre de notre proposition, cette nouvelle littératie des données doit être constitutive de la « justice des données » en tant que réponse aux conséquences sociétales des technologies et des services basés sur les données qui engendrent inégalités et discriminations au sein des logiques automatiques et algorithmiques (Dencik, 2019 ; 2022). Il s'agit de favoriser une prise de conscience sur les pratiques de données dont les incidences se répercutent sur l'expérience² de vie des personnes *queer*.

Constats

Activisme des données et violence des données

- 6 Comme précisé, les études critiques sur les données enquêtent sur le « pouvoir » qui imprègne les pratiques des données. Ce pouvoir est ambivalent (Hepp *et al.*, 2022) : son attrait découlant du fait que les données numériques permettent la production de savoirs qui nous renseignent sur la société et les processus sociaux, ainsi que sur les problèmes que ces mêmes données engendrent, notamment la surveillance de masse et

les reproductions des inégalités amenant à des pratiques discriminatoires (Cremonini, 2023 ; Leurs et Sheperd, 2017). En ce sens, les études critiques doivent répondre à cette ambivalence et développer une recherche critique crédible de ce pouvoir, en particulier la capacité de déployer des tactiques de résistance au détournement des données personnelles pour les réemployer et répondre à des objectifs plus personnels et sociaux.

- 7 De ces visions s'est institué l'activisme des données, « un outil heuristique et polysémique employé pour réfléchir politiquement au *big data* selon la perspective des utilisateurs et des citoyens, en analysant en particulier l'évolution contemporaine de l'activisme vis-à-vis de la datafication³ » (Milan et Velden, 2016). Il se situe aux frontières de l'activisme écologique, de la remise en cause du processus de la *datafication* et de ses conséquences socio-politiques, notamment du fait qu'il redéfinit et restructure à la fois les environnements numériques et physiques pour adapter le travail aux exigences de l'économie en réseau (Ortiz-Freuler, 2023). Il doit permettre « de réfléchir sur l'ubiquité des communications numériques et sur les dynamiques de médiation ayant cours dans la société des plateformes » ainsi que sur la « nature de l'activisme contemporain et de l'évolution de ses configurations médiatiques et de protestation⁴ » (Milan et Velden, 2016).
- 8 Il s'inscrit dans une logique politique nécessitant de prendre part aux initiatives liées aux données et de résister aux stratégies de collectes, en particulier des données personnelles, pour éviter que ne se reproduisent des cas comme l'affaire Cambridge Analytica où les données de 87 millions d'utilisateurs de Facebook avaient été siphonnées. En cela, les activistes des données se posent en intermédiaires pour soutenir l'agentivité des « publics transformés en données » (*datafied publics*) et créer des imaginaires alternatifs à la *datafication* (Miloni et Papa, 2022).
- 9 Au sein des luttes menées par cet activisme, la question des problématiques liées aux expériences de vie de groupes marginalisés est centrale. En particulier dans les données où le terme de « violence des données » (*data violence*) désigne pour Hoffmann (2021) les violences matérielles, symboliques, ainsi que toutes autres violences infligées par et à travers les technologies des données et leurs fournisseurs. Ces violences sont omniprésentes, fluides (Keyes, 2019) et se diffusent à grande échelle sous trois formes.
- 10 Tout d'abord, par l'invisibilisation des données *queer* qui a été surtout mise en avant dans les travaux portant sur les recensements des populations. Pour Reynolds (2001), ces derniers sont partiels, loin de refléter une supposée neutralité des méthodes de recueil des données (Browne, 2010). Neutralité qui cache en fait un « catalogue de choix » jugé trop réducteur pour prendre en compte la complexité des vies *queer* (Ruberg et Ruelos, 2020 ; Guyan, 2022). Les recensements doivent ainsi nous forcer à nous questionner sur les méthodes de recueil des données ainsi que sur les résultats qui en sont tirés (Browne, 2010).
- 11 C'est toute la problématique de « l'inclusion » qui est discutée. Située au cœur de la conception des jeux de données, elle est considérée comme une sorte de technologie de gouvernance qui « transforme les étrangers en sujets, ceux qui sont inclus sont également prêts à consentir aux conditions de l'inclusion⁵ » (Ahmed, 2012, p. 163). Tomasev *et al.* (2021) ont en démontré les limites, soulignant notamment la difficulté d'observabilité des religions, des handicaps et des types de classes sociales des publics convertis en données. Ainsi l'on retrouve, au-delà des recensements, des problématiques concernant la non-reconnaissance d'un genre (Keyes, 2018), en

particulier dans le milieu médical, qui entraîne des difficultés de traitement pour les personnes transgenres, non-binaires ou *gender queer*⁶ (Scandurra et al., 2019 ; Alessandrin et Mediani, 2022), invisibilisant leur existence dans les bases de données. Par exemple, en janvier 2025, les agences de santé américaines ont retiré les pages web intégrant les statistiques de santé des personnes *queer*⁷ (notamment toute information liée au VIH), amplifiant les risques d'une crise sanitaire.

- 12 Ensuite, la deuxième forme, l'essentialisation, a été étudiée dans l'analyse des jeux de données comme celle controversée de Wang et Kosinski (2018) pour qui, sur la base d'un algorithme qui analysait les traits faciaux des personnes, il était supposément possible de déterminer l'orientation sexuelle de ces dernières. Cette thématique a pris une nouvelle ampleur avec la diffusion des IA génératives : en s'entraînant sur de nombreux jeux de données, ces dernières ont produit des profils *queer* caricaturaux. Un exemple récent est celui de *Wired* qui a employé le logiciel *Midjourney* pour générer des images de personnes *queer*⁸ : ces dernières étaient très sexualisées et représentées avec des cheveux colorés, une peau souvent blanche ou encore une carrure masculine dans le cas des femmes transgenres. Ceci est à prendre en compte dans un contexte où le manque de données originales (Jones, 2024) va pousser les organisations à entraîner les IA sur des données générées par ces dernières, renforçant les problématiques évoquées ci-dessus.
- 13 Enfin, la dernière forme est la stigmatisation et est la plus difficile à référencer du fait de son caractère insidieux. Keyes (2019) soulignait que les normes et les formes de violence existent en tant que système qui nous force à nous conformer à un moule (en termes de personnalités et de comportements à adopter) : dévoiler les caractéristiques des personnes *queer* les expose ainsi à des périls, en particulier dans les pays où la non-hétérosexualité et la non-conformité liée au genre sont punissables de mort. La stigmatisation se dévoile ici en particulier lors d'un mauvais usage des données des personnes *queer* librement accessibles comme dans le cas des applications de rencontres. Dans plusieurs pays où des applications comme *Grindr*⁹ sont autorisées, plusieurs faits de surveillance¹⁰, de trafic de données (Kokas, 2023) ou encore de guet-apens¹¹ ont été recensés, posant plus fortement la question de la sécurité des dites personnes *queer* et de leurs données.

Les études et littératies des données

- 14 Vis-à-vis de problématiques similaires, la littératie des données a été considérée comme limitée en termes de compétences à développer, de thématiques à analyser ou encore de dénominations inadéquates, amenant à des évolutions rhizomiques de celle-ci pour compenser ces manques (Verdi, 2023 ; 2024). En effet, la littératie des données a fait l'objet dès 2010 d'appropriations par de nombreux auteurs, chaque proposition visant à dépasser ce cadre limité de la vision généraliste pour pouvoir développer des connaissances et des compétences sur des sujets non envisagés originellement. Il existe actuellement plus d'une trentaine de littératies des données différentes (Verdi, 2023) qui traitent par exemple de domaines d'études comme la santé (*healthcare data literacy*), de données spécifiques (ex. : *big data literacy*, *open data literacy*, *linked open data literacy*), de l'enseignement (ex. : *educational data literacy*) ou encore de compétences intellectuelles comme le développement d'un recul critique dans l'usage des données (ex. : *critical data literacy*).

- 15 Si les études critiques des données et les littératies *queer* abordent les problématiques sociétales liées au genre et au sexe (Edward et Greenough, 2019 ; Govender et Andrews, 2021 ; Miller, 2015 ; Potter, 2022 ; Sandretto, 2018 ; Sykes, 2011 ; Threlkeld, 2011), nous constatons néanmoins l'absence d'une littérature traitant spécifiquement des questions *queer* dans les données. C'est pourquoi nous souhaitons ici combler ce manque.

Usage du *queer*

- 16 Pour ce faire, nous recourons au concept polysémique du *queer* pour englober le plus grand nombre d'identités et permettre de créer des données *queer* ou de « queeriser » celles existantes. Les données *queer* (*queer data*) désignent toute recherche sur les conséquences socioculturelles des données (au sens des réalités qu'elles embarquent, qui les prédéfinissent) pour déterminer la corrélation entre les expériences de vie des personnes *queer* et leur transcription au cœur des standards des données (Duncan, 2021 ; Guyan, 2022). Pour Ruberg et Ruelos (2020), il s'agit de « remettre en question et résister aux attentes et aux normes qui entourent actuellement les données et la datafication¹² ».
- 17 La *queerisation* prend donc la forme d'une méthode « anti assimilationniste » permettant d'exploiter « les ressources de la marge » (Bourcier, 2018) afin de se soustraire aux visions binaires sur l'identité et corriger les jeux de données, comme dans le cas des archives *queer*. Nous pouvons ici citer le réseau Big Tata¹³ qui s'appuie sur l'Homosaurus¹⁴, un thésaurus international de données liées portant « sur les personnes, la culture et les enjeux LGBTQ+ », pour décrire ses contenus au sein de son catalogue collectif.

Les racines du *queer*

- 18 Provenant de la racine indo-européenne « *twerkw* » signifiant « à travers » (Sedgwick, 1998), le terme *queer* désigne toute personne dont le comportement s'éloigne des normes socialement acceptables. Il a ensuite été utilisé comme insulte à l'encontre des personnes déviant de la norme sexuelle (Eribon, 2003). Il a été redéfini par Sedgwick (1998) comme « le maillage ouvert de possibilités, de lacunes, de chevauchements, de dissonances et de résonances, de laps et d'excès de signification, quand les éléments constitutifs du genre de quiconque, de la sexualité de quiconque, ne sont pas faits (ou ne peuvent pas être faits) pour signifier de manière monolithique » (traduction de Zielinska, 2022). Selon Damian-Gaillard et Vörös (2023), les études *queer* « croisent la lecture de Michel Foucault sur le dispositif de sexualité avec celles d'Adrienne Rich sur la contrainte à l'hétérosexualité et de Monique Wittig sur l'hétérosexualité comme système politique pour penser l'hétéronormativité comme un mécanisme de hiérarchisation de sexes et des sexualités fondé sur l'idée de nature ». Cette hétéronormativité naturalise l'hétérosexualité, l'hétérosexisme, l'hétéropatriarcat et les normes cisgenres pour en faire le point de référence des pratiques sexuelles (Govender et Andrews, 2021). Elle ne prend pas en compte celles et ceux qui existent en dehors de cette norme et ignore leurs revendications (Reynolds, 2001). Mieli *et al.* (2017) emploient, eux, le terme d'« éducastration » pour souligner l'assignation automatique d'une hétérosexualité conforme au genre attribué à la naissance.

19 Le *queer* plonge ses racines dans les théories féministes (Rosenberg, 2008) qui se retrouvent dans la « théorie queer », terme employé par plusieurs auteurs comme De Lauretis, Morris et Anzaldúa, qui attire l’attention sur « le travail conceptuel et spéculatif qu’implique la production du discours et la nécessité d’un travail critique qui consiste à déconstruire nos propres discours et nos silences construits » (De Lauretis, 2023, p. 91). En effet, il s’agit, comme le rappelle Perreau (2018, p. 116), d’une démarche critique, stratégique, attentive aux déplacements, aux zones de friction ainsi qu’aux échecs de la norme. Une critique qui « problématise et politise non seulement le corps, mais aussi [...] le savoir et la production de vérité, bref les rapports savoir-pouvoir » (Bourcier, 2018, p. 147). Une synthèse des différents courants de cette théorie a été réalisée par Bernini (2024) (Tableau 1).

Tableau . Courants des théories *queer*, d’après Bernini (2024)

Freudo-marxisme révolutionnaire	Constructivisme radical	Théories antisociales
« Approche qui associe les intentions révolutionnaires du marxisme au potentiel de libération présent dans les théories freudiennes, opérant une double critique dirigée conjointement contre la société capitaliste et la répression sexuelle ». Elle « pense l’action du pouvoir sur la sexualité essentiellement sous la forme d’une répression et envisage la sexualité avant tout comme un désir sexuel refoulé qui ne demande qu’à être libéré » (p. 106-107).	« Aucun ancrage extérieur au dispositif de sexualité (ni le désir, ni le sexe) n’est accessible au sujet sexuel, et la résistance ne devient possible qu’en s’appuyant sur les normes dictées par ce même dispositif » (p. 132). Les théories queer constructionnistes radicales ont de commun avec celles freudo-marxistes une attitude antagoniste, mais aussi optimiste et confiante dans les possibilités de transformation de l’action politique (p. 133).	Au contraire des précédentes théories qui prônent une résistance au dispositif de sexualité par la construction de nouvelles subjectivités et de nouvelles communautés de reconnaissance, les théories antisociales queer insistent sur l’irruption de la pulsion sexuelle comme moteur de la rupture du lien social et de la suppression du sujet dans la jouissance. La négativité liée aux pratiques sexuelles devant être remises en avant (p. 134-136).

20 Cette synthèse montre qu’il n’existe donc pas une, mais des théories *queer* qui se subdivisent en plusieurs courants. Or, comme le rappelle la synthèse de Back (2024), les théories *queer* sont issues en majorité d’auteur·rices *queer* blanc·he·s, incorporant des angles morts dans les analyses et invisibilisant nombre de corpus *queer* racisés. À défaut de proposer de nouveaux cadres théoriques, à l’instar de Bumaye (2024), nous prendrons en compte les problématiques inhérentes ici soulignées et intégrerons une approche intersectionnelle que nous définirons plus tard.

Déconstruire les normes

21 L’objectif du *queer* n’est pas un rejet, mais un décentrage des régimes de normalité, tout particulièrement des standards hétérosexuels et cisnormés, pour englober l’ensemble du spectre des identités (Duncan, 2021). Pour Niedergang (2023, p. 31), la critique *queer* s’adresse non à la normativité, mais à la normalisation, à savoir « la tendance à juger

des comportements sexuels, les identités et les désirs, à partir d'un seul ensemble de normes hégémoniques et dominantes issu d'une seule matrice ». Dans ce cadre, la critique de la normalisation n'implique pas l'anormativité, mais une normativité *queer* : c'est faire « dérailler » les normes afin de remettre en cause la « matrice normative » (p. 29).

- 22 Cette matrice normative hétérocisnormée assimile ainsi comportements et catégories sexuelles, mettant au centre des réflexions *queer* la question du genre. Le terme de genre est originellement défini par Money (1955) qui a décorrélé les « rôles de genre » des sexes biologiques. Il a ensuite relevé des rapports de pouvoir (Butler, 1997), étant constitutif des rapports sociaux en raison des caractéristiques spécifiques attribuées aux différents sexes (Scott et Varikas, 1988). En effet, les « personnes » ne deviennent intelligibles que si elles ont « pris un genre » (*becoming gendered*) selon les critères distinctifs de l'intelligibilité de genre (Butler, 2006, p. 83) et ce genre est corrélé au sexe biologique.
- 23 Le lien entre genre et sexe est donc central. Delphy (2001) affirme, contre ce qu'elle nomme la « théorie cognitiviste » (p. 250), que c'est le genre qui a une antécédence sur le sexe, ce qui implique de penser le sexe comme un « marqueur de la division sociale » (p. 252) amenant à des visions catégorielles de type binaire (homme-femme), immuable (on ne change pas de catégorie une fois celle-ci attribuée) et physiologique (assignation basée sur des caractéristiques externes comme les organes génitaux) (Keyes, 2018). Pour De Lauretis (2023), le genre est une « technologie du sexe ». En tant que « représentation et auto-représentation », il est le produit de technologies sociales : en effet, le genre, comme la sexualité, est « un ensemble d'effets produits dans les corps, les comportements et les relations sociales » grâce au déploiement d'une « technologie politique complexe ».
- 24 L'objectif des théories féministes est ici de dépasser ce « système sexe/genre » (De Lauretis *et al.*, 2023, p. 44). C'est, d'une part, démontrer que comportements et catégories sexuels ne coïncident jamais parfaitement et que les sujets se constituent à travers cette non-coïncidence même (Perreau, 2018). D'autre part, rappeler que le genre n'est pas « le miroir du sexe » (Cervulle, 2023).

Identités *queer*

- 25 Le *queer* est ainsi « non seulement la réappropriation d'une interpellation qui, permet de déplacer les catégories, méthodes et champs du savoir légitime, mais il porte également une attention particulière à toutes sortes de traversées », qu'il s'agisse des frontières catégorielles de la nation, de la race ou de l'ethnicité (Perreau 2018). Il n'est pas une catégorie, mais un « mouvement libre dans l'espace résiduel entre les catégories¹⁵ » (Schram, 2019). Il ne peut donc être pleinement possédé, mais seulement « redéployé, tordu, queerisé à partir d'un précédent usage et orienté vers des objectifs politiques urgents et croissants¹⁶ » (Butler, 2011) invitant ainsi à une nouvelle signification de l'identité.
- 26 Du latin *identitas* (Oudin, 2012), l'identité est la source de la signification et de l'expérience des individus (Castells, 2009, p. 6). Elle est à la fois le résultat d'un processus collectif, historique, social, politique (la construction du peuple, d'une nation) et un phénomène qui touche à l'âme, à l'intime, au plus personnel (Ollivier, 2007, p. 56). Elle revêt la « dimension symbolique d'une forme et d'une représentation

dotée pouvant faire l'objet d'une interprétation, d'une diffusion et d'une acquisition » (Lamizet, 2002), car intégrée dans des pratiques de communication où elle devient « un système dynamique de relations et de représentations [...] qui implique la reconnaissance mutuelle du Moi et de l'Autre » (Dymytrova, 2011, p. 21).

- 27 Dans un cadre *queer*, il s'agit d'embrasser une multitude d'identités et de pratiques de soi (Cervulle et Quemener, 2021). L'identité n'est ni figée ni unique, car en réalité « chaque individu possède une pluralité d'identités » qui se construisent dans les interactions sociales (Mellini, 2009, p. 6-7). Il ne s'agit plus de parler « du » genre, mais « des » genres ; le terme « genres » pouvant ainsi être utilisé pour les identités, à savoir les rôles sociaux sexués performés par les individus (Marignier, 2016, p. 38).
- 28 Le *queer* s'oppose ainsi souvent à l'identité civile, c'est-à-dire à l'ensemble des caractéristiques attribuées par l'État qui désigne la qualité de citoyen comme le nom, l'âge, le sexe, etc. Tout particulièrement dans un cadre transgenre où l'identité de genre fait l'objet de combats militants, notamment dans la facilitation de changement d'état civil. Recourir au *queer* s'avère ici être « une école de responsabilité dans la prise de risque d'une invention - ou d'une réinvention — permanente de soi et des autres » (Bordas, 2004).
- 29 Découlant de cela, un autre point important des théories féministes qui se retrouve dans le *queer* est le lien entre genre et technologie (De Lauretis, 1987). Ceux-ci s'influencent et se définissent mutuellement (Landström, 2007) : le genre est médié par la technologie et les technologies du genre qui, en tant qu'ensemble de pratiques sociales et discursives, produisent le sujet comme « genré » (Meidani et Alessandrin, 2019). Dans ce contexte, le genre ne peut pas être « performé indépendamment des structures classificatoires sociotechniques que les humains intègrent aux ordinateurs et que les ordinateurs attendent des humains¹⁷ » (Simeoni *et al.*, 2024). En particulier dans les pratiques de *datafication* où les données produites concourent à la définition des identités. Or, nous avons pu voir les limites de ces pratiques quand elles concernent les expériences de vie *queer*.

Proposition d'une littératie des données *queer*

- 30 Nous proposons de poser les bases d'une littératie des données *queer*, définie comme une littératie des données centrée sur une quintuple approche (critique, mécanologique, *queer*, intersectionnelle et éthique) permettant de (1) prendre une part active, politique, aux débats sur les aspects éthiques, idéologiques et de pouvoir des données qui influencent les vies *queer*, en particulier dans un contexte de *datafication*¹⁸, et de (2) proposer des méthodes de conception, d'analyse et de représentation de données *queer*.

L'ancrage littéracique

- 31 La littératie est la qualité ou l'état d'un lettré (*literate*), à savoir celui qui sait lire et écrire¹⁹. Elle désigne un ensemble de pratiques et de compétences qui permettent de construire des savoirs, soulignant les effets de la lecture et de l'écriture sur la cognition humaine (Olson et Lejosne, 2006). Du fait qu'elle est l'« ensemble des praxis et représentations liées à l'écrit » (Privat, 2010), la littératie structure la pensée et façonne nos interactions avec le monde, faisant d'elle une *Weltanschauung*, une conception du

monde couplée à la manière dont l'humain conçoit ce monde (Jung, 1996). Or, parler de « la » littératie est incorrect du fait qu'il n'existe que des « constellations de formes et de degrés de littératie²⁰ » (Sledd, 1988), rejoignant Olson (1998) qui précisait qu'il existait non pas une, mais des cultures de l'écrit « fort différentes les unes des autres ». C'est pourquoi nous parlerons toujours « d'une » littératie des données *queer*.

Les pédagogies et littératies *queer*

- 32 Au même titre que la littératie qui incorpore la transmission des savoirs écrits par l'instruction et l'apprentissage (Norris et Phillips, 2003), nous intégrons dans notre proposition diverses méthodes issues des pédagogies *queer*. Ces dernières émergent dans les années 1990 (Bryson et De Castell, 1993 ; Britzman, 1995 ; Luhmann, 1998) et, inspirées des théories *queer*, cherchent à contribuer au développement de pratiques éducatives qui analysent la mobilité et la fluidité de la société tout en soutenant l'idée que se baser sur un modèle éducatif unique finit par aliéner, voire discriminer, certains individus. Il ne s'agit pas de chercher la « bonne » méthode ou de se poser les « bonnes » questions, mais plutôt de pouvoir remettre en cause les pratiques éducatives traditionnelles, perçues comme naturelles, ainsi que les notions d'égalité et d'acceptation (Nemi Neto, 2018). En résumé, les pédagogies *queer* s'orientent autour de deux modalités centrales : l'anti-normativité et la subversion sur laquelle nous reviendrons (Poirier-Saumure, 2024).
- 33 Les littératies *queer* sont, elles, définies par Wargo (2022) comme les pratiques discursives, les réalités vécues et les processus divers de communication que les personnes et communautés *queer* emploient dans leur résistance aux régimes normatifs du genre et de la sexualité. Elles ont émergé avec Miller (2015) et Helmer (2016) où le *queer*, au-delà de faciliter une nouvelle lecture des expressions (a)genrées et (a)sexuelles, revêt la forme d'une conscience politique qui offre la possibilité d'une autodétermination des identités. En commun avec les pédagogies *queers*, elles se centrent sur les manières d'opérer des changements des pratiques ayant lieu au sein des salles de classe pour éviter les discriminations et encourager l'autodétermination des identités des étudiants (Van Horn, 2022).

La donnée comme système

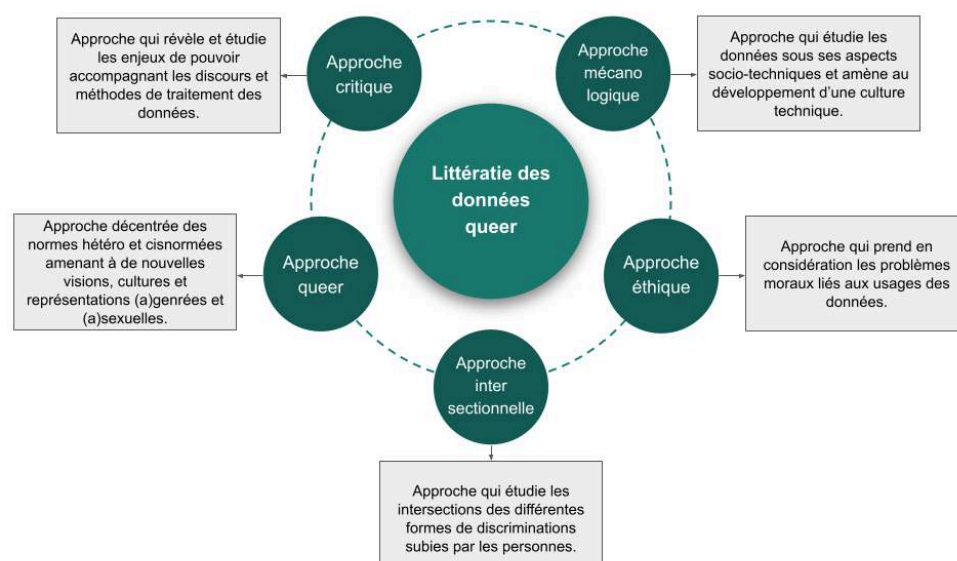
- 34 Du fait notamment de sa polysémie (Hjørland, 2018) et de son rôle dans un système technique complexe (comme les mécanismes qui les relient aux algorithmes et aux IA), les données sont difficiles à comprendre et à enseigner (Verdi, 2024). Pour faciliter cela, nous considérons que les données doivent être comprises comme un système²¹ et non comme des entités séparées. Parler des données, en particulier dans un contexte de formation, c'est nécessairement mentionner toute la structure matérielle qui les porte, les logiciels qui les traitent, ou encore les algorithmes et les IA qu'elles alimentent, pour permettre aux futur.es lettré.es d'avoir une vision complète de ce que signifie concevoir et traiter des données qui, au contraire de nombreux discours (Verdi, 2024), sont loin d'être neutres et immatérielles. C'est donc percevoir la donnée comme un système qui se détache « des » données pour subsumer un écosystème numérique plus vaste où les problématiques *queer* se retrouvent (Tacheva et Ramasubramanian, 2023).

Les ancrages

- 35 En raison de la symbiose entre les éléments précités (notamment les données avec les IA), nous soutenons une symbiose entre les littératies afin d'enrayer la création de littératies souvent éloignées les unes des autres et simplement centrées autour des concepts qu'elles abordent. Notre vision est systémique et distribuée au sens d'une littératie distribuée (Verdi et Le Deuff, 2020) : la littératie des données *queer* ne peut se concevoir comme une littératie autarcique, mais, à l'inverse, symbiotique avec (1) les littératies des algorithmes et des IA, rappelant ici la nécessaire prise en compte des logiques sociotechniques dans lesquelles s'insèrent les données, (2) les littératies informationnelles et des données pour rappeler leurs liens conceptuels et (3) les littératies critiques et *queer* en intégrant différentes postures intellectuelles et méthodologiques pour concevoir et analyser les données. Elle s'inscrit enfin dans (4) la littératie numérique conceptualisée par Vitali-Rosati (2024) à travers ses trois principes : la conscience de la multiplicité des modèles (comprendre quel est le modèle qui fonde une technologie particulière), la recherche de la complexité (la matérialité porte le sens) et la maîtrise de l'activité (chercher activement des solutions).
- 36 Ces ancrages doivent nous permettre d'une part de rapprocher la littératie des données *queer* de la vision de Keyes (2019) qui proposait le développement d'une science des données radicale offrant aux individus diverses manières d'être eux-mêmes et ne punissant pas ceux qui ne « participent pas au système ». D'autre part, de s'inscrire dans la vision de Wang (2024) d'une *datafication* située qui « localise, fait émerger et intègre les contextes historiques, sociaux, politiques et de valeurs des agents instruits — qu'ils soient humains, machiniques ou cybernétiques — dans les procédures de représentations et opérationnelles de la *datafication*²² ».

Les cinq approches

- 37 La littératie des données *queer* se décline selon cinq approches interdépendantes que nous allons maintenant présenter (Figure 2).

Figure . Approches de la littératie des données *queer*

Source : Ugo Verdi

L'approche critique

- 38 Inspirée des études et littératies critiques, notre littératie doit permettre d'explorer les enjeux de pouvoir à l'œuvre dans les données, de révéler plus précisément les idéologies et structures de contrôle et de domination. Nous reprenons à notre compte la vision de Brand et Sander (2020) qui présentaient la littératie critique des données comme une compétence permettant de traiter la *datafication* de manière critique en réfléchissant aux implications sociétales du traitement des données et en mettant en pratique cette compréhension. Une conception et une interprétation d'un autre « monde des données » (Cardon, 2012) qui doit intégrer la subversion comme « bouleversement de notre capacité à lire le monde normatif pour ensuite révéler de nouvelles possibilités de lecture émancipatoires » (Poirier-Saumure, 2024). En reprenant l'exemple des recensements, c'est ici déroger aux pratiques généralistes de catégorisation pour élargir le spectre des possibilités.
- 39 C'est comprendre que, comme le genre, les données sont des *capta* (Lanigan, 1994 ; Latour, 2007 ; Drucker, 2011). Qu'en tant que matériaux produits par l'abstraction du monde en catégories, mesures et autres formes de représentations qui forment le socle des blocs constitutifs à l'élaboration de l'information et du savoir (Kitchin, 2014), elles sont une vision subjective des réalités, soulignant à nouveau leur aspect non brut (Gitelman, 2013). Elles ne parlent pas d'elles-mêmes : seules une narration et une mise en relation avec d'autres données permettent de leur attribuer un sens (Bachimont, 2020 ; Floridi, 2009 ; Leleu-Merviel, 2016 ; Verlaet, 2024). Un sens médié par les dispositifs numériques qui conditionnent les données et le message qui en sera traduit comme dans le cas de la datavisualisation.
- 40 C'est en outre, s'inscrire dans une logique de dépassement de la « rhétorique de l'immatérialité » (Vitali-Rosati, 2024, p. 83) en remettant la matérialité des données au centre des discours. Notre littératie hérite ici des traditions du matérialisme numérique évoquées par Casemajor (2015) qui font des données numériques des inscriptions

physiques (Kirschenbaum, Ovenden et Redwine, 2010) ayant des effets matériels, comme le coût écologique de leur gestion (Flipo, 2020 ; Crawford, 2022).

- 41 Les lettrés des données *queer* doivent être ici en mesure de déterminer les aspects politiques, sociaux et matériellement construits des données (Irgens *et al.*, 2020), d'avoir conscience du rôle et du pouvoir des discours sur les données au sein desquels l'expérience se forme (Anderson et Irvine, 1993). Les techno-discours sont ici particulièrement visés du fait qu'ils limitent les techno-imaginaires (Balandier, 2001) en empêchant toute remise en question du phénomène technique (Janicaud, 1985) et enfermant leurs récepteurs dans un impensé technique qui installe les technologies dans la posture de l'évidence (Robert, 2020).
- 42 Parmi ces impensés, nous pouvons citer les données perçues comme un « or noir » permettant d'alimenter des tableaux de bord indiquant les meilleures décisions à prendre. Un point qui a été nuancé avec les critiques liées au *big data* (Crawford, 2022 ; Cremonini, 2023). C'est ainsi adopter une posture de « désenchantement » (Verdi, 2024) qui dépasse les visions pharmaniques²³ sur les données et contrecarre les discours « enchantés » sur les technologies (Citton, 2017).
- 43 Ceci nécessite une analyse des textes et des discours, intégrée dans une logique globale de développement de l'hétéroglossie, à savoir « un autre discours dans une autre langue, qui sert à exprimer les intentions de l'auteur, mais de façon réfractée » (Bakhtine, Holquist et Emerson, 1982 ; Street, 2006), où le mot est compris une relation de pouvoir, d'identité, de différence et d'accès à la connaissance, à des compétences, des outils et des ressources (Janks, 2013). Et par-delà, de pouvoir négocier les différentes lectures des données (Irgens *et al.*, 2020) qui nécessitent du lettré.e une participation politique (Stornaiuolo, 2020), rejoignant la pensée de Paulo Freire dont héritent les études et littératies critiques sur les données (Hautea, Dasgupta et Mako Hill, 2017).

L'approche mécanologique

- 44 Nous intégrons également une approche mécanologique telle que reconceptualisée par Le Deuff et Roumanos (2022) comme « celle qui permet d'ouvrir le dispositif non seulement au sens technique, mais également au sens de l'étude de la machine qui produit des effets et qui est pensée avant même sa concrétisation ». Au sein d'une telle approche doivent être développées une conscience des limites et biais des dispositifs technologiques (Crawford, 2013 ; O'Neil, 2017, 2022 ; Tripaldi et Morelli, 2024) ainsi qu'une capacité de les contourner, notamment à travers des méthodes de surveillance (Bridges, 2021 ; Wolff, 2022).
- 45 C'est développer une culture technique nécessaire à la compréhension des données et de leurs éléments connexes, définie comme « la possession des connaissances et du savoir-faire susceptibles de donner une maîtrise personnelle de notre environnement et un contrôle sur l'activité de ceux à qui l'on doit avoir recours » (Ellul, 1988, p. 170) afin d'atteindre un équilibre entre un état de « minorité » (l'objet technique vu avant tout comme objet d'usage et où le savoir technique est implicite/non réfléchi/coutumier) et un état de « majorité » (prise de conscience et opération réfléchie de l'adulte libre qui a à sa disposition les moyens de la connaissance rationnelle élaborée par les sciences) (Simondon, 2024, p. 125).

- 46 La culture technique nécessite de « bien connaître le milieu dans lequel on vit » sous peine d'une double aliénation : méconnaissance de l'environnement et dépendance à ceux qui ont la connaissance (Ellul, 1988, p. 170). Et ainsi d'adopter une « posture technologique » où l'on « ne se préoccupe pas seulement de l'usage d'un être technique, mais de la corrélation des êtres techniques les uns par rapport aux autres » (Simondon, 2024, p. 216).
- 47 Pour illustrer les points suscités, nous évoquons l'article à venir de Le Deuff (2026) qui présente l'usage des petits objets connectés ludiques (POCL), à savoir « des dispositifs électroniques compacts et inventifs, conçus pour matérialiser de manière tangible et ludique des flux de données numériques ». Il s'agit ici de représenter et comprendre des données souvent abstraites (ex. : la qualité de l'air) à travers des objets tangibles (ex. : un parapluie s'ouvrant en cas de pluie) dont la conception technique sera documentée pour pouvoir être reproduite.
- 48 La culture technique doit également révéler les « bugs » (Vitali-Rosati, 2024) pour adapter les décisions à prendre. *Garbage In, Garbage Out* : si les données sont mal configurées, les résultats seront médiocres, amenant à l'émergence d'informations biaisées, voire fausses. Ce qui est amplifié aujourd'hui avec les IA du fait qu'elles rendent simultanément « la catégorisation et le tri plus dynamiques et plus probabilistes » et intègrent ces capacités au sein « des systèmes quotidiens -- c'est-à-dire dans le cadre des infrastructures (numériques)²⁴ » (Taylor *et al.*, 2025) ; nous l'avions mentionné, en introduction, dans le cas des images *queer* générées par Midjourney.
- 49 C'est en ce sens une cassure contrevenant aux logiques informationnelles qui nécessite un re-travail des représentations, car ces dernières ne font pas que re-présenter une réalité déjà existante : elles sont aussi le seuil de nouveaux futurs potentiels (Gossett, Stanley et Burton, 2017). Philip, Olivares-Pasillas et Roca (2016) ont mis en avant le fait que la datavisualisation, en tant que lieu potentiel de micro-agressions, est perçue comme un « artefact » pouvant reproduire les inégalités sociales et les hiérarchies sociétales. Afin de les éviter, les auteurs l'ont réemployée pour souligner les biais raciaux présents dans les jeux de données et inviter à un examen collectif des postulats idéologiques. C'est, en somme, avoir conscience que les données, en tant qu'objet technique, possèdent un « contenu politique », car constituant « des éléments actifs d'organisation des relations des hommes entre eux et avec leur environnement » (Akrich, 1987).

L'approche queer

- 50 La troisième approche est celle du *queer*. Au-delà de la nécessaire analyse des discours, d'autres thématiques doivent être évoquées. En premier lieu, celle des représentations et des identités pour la constitution et l'analyse des données. Comme toute littérature *queer*, la nôtre n'assigne pas de facto un genre ou une sexualité aux personnes : elle invite ces dernières à s'autodéfinir à travers leur propre vocabulaire et à prendre en compte le fait que cette autodétermination peut être temporaire, fluide et complexe (Duncan, 2021), et ce, à l'aide d'outils comme le *Queer Diamond* (Simeoni *et al.*, 2024) qui permet de visualiser le spectre du genre et donc d'aider les personnes à mieux déterminer leur identité de genre et les genres auxquels elles sont attirées.

- 51 Nous nous inspirons ici de la théorie émique de Browne (2010), tirée des travaux du philosophe Kenneth Pike, qui désigne une méthodologie située prenant appui sur le point de vue de l'objet étudié. À cette fin, l'autoethnographie, en tant que méthode de recherche à l'intersection de l'autobiographie et l'ethnographie qui se base sur l'expérience personnelle (« auto ») pour décrire et interpréter (« graphie ») les textes, expériences, croyances et pratiques culturelles (« ethno ») (Adams, Ellis et Holman Jones, 2017), permet d'apporter des détails de la vie quotidienne invisibles aux personnes non concernées afin de remettre en question les méthodes traditionnelles de recueil des informations et lutter contre les stéréotypes.
- 52 Dans le cas de l'autoethnographie *queer*, l'objectif n'est ainsi pas tant de documenter les vies *queer* que de modifier les pensées conventionnelles sur le genre et la sexualité (Wilson, 2019). Plusieurs auteurs s'en sont emparés pour constituer leur identité, vue comme une « contre-archive vécue » (Jahjah, 2024). Bouqentar (2021), en tant que chercheuse *queer* et arabe, mixe cette approche avec celle de « l'identification comme devenir » ainsi que de la notion d'identification de Hall (1997) et de Silva (2010), qui renvoie selon elle à « un positionnement mouvant, fluide, que le sujet assume, occupe, dans des contextes, moments et situations donnés » afin de déployer un cadre anti-essentialiste, lui permettant de « partir du trivial qui compose la quotidienneté » et de mettre en lumière « les différentes oppressions qui émergent comme effet des identification *queer* et arabe dans un contexte migratoire ».
- 53 Jahjah (2022 ; 2024) étudie lui le racisme sexuel et la fétichisation raciale à partir de son expérience de personne racialisée. Il déploie à ce titre une pluralité de « positionnalités interventionnistes », en particulier à travers son personnage de « Lady Booba » sur l'application de rencontres *Grindr*, usant de filtres et alternant les photos de profil dans une stratégie de désidentification. En tant qu'« énoncé fabulatoire », son personnage répond « aux tentatives de naturalisation de catégories historiquement forgées ».
- 54 C'est ainsi une prise en compte des identités qui s'expriment différemment selon l'espace-temps dans lequel elles sont situées. Comme l'a précisé Foucault (1984), l'espace est fondamental dans l'exercice du pouvoir et tout particulièrement l'espace public. Si Habermas (1988) précise que ce dernier doit permettre d'amener à l'expression d'un intérêt général, Borghi (2012) rappelle qu'il est « conçu, géré et modelé sur la base d'une conception dualiste rigide : homme-femme, licite-illicite, homosexuel-hétérosexuel ». Elle ajoute que la sexualité « touche les espaces du quotidien » et « revêt une portée géopolitique significative » (Borghi, 2014). C'est pourquoi une volonté de « déterritorialisation » de l'hétérosexualité s'est faite jour afin d'intervenir dans les « dispositifs biotechnologiques de production de subjectivité sexuelle » (Preciado, 2003).
- 55 Zwer (2024) a ainsi proposé une « contre géographie » féministe : elle y décrit l'espace comme un lieu de luttes et de discrimination, rejoignant ici les travaux d'Alessandrin et Dagorn (2023) sur l'espace urbain présenté comme hostile aux personnes *queer*, aux personnes racisées et aux femmes. La production de l'espace étant tributaire « des relations mouvantes du pouvoir entre actrices et acteurs à tous les échelles, dans une "géométrie du pouvoir" » (Zwer, 2024, p. 128), le langage lui-même qui imprègne les cartes doit être déconstruit, amenant à la création de « contre cartes ». Ces dernières, instruments centraux de la contre géographie, doivent se faire à partir « de la dimension vécue, ressentie, imaginaire de l'espace » (p. 168), recueillie auprès des personnes qui habitent ces espaces, à l'instar du site *Queering the Map* qui permet aux

internautes de décrire en ligne sur une carte du monde leur expérience *queer* des espaces physiques.

- 56 Au sein de ces espaces, le corps (en tant que vecteur d'identité, d'action, de discours, du genre et de la sexualité) (Adam, 2018, p. 49) est de fait orienté par des forces politiques (Butler, 1997) : il est l'effet d'un rapport de pouvoir « au sens où il est façonné, discipliné par ce rapport, qui renvoie à un système de domination articulé à l'hétérosexualité obligatoire » (Dorlin, 2008). Combattre cette normalisation nécessite de proposer « un corps agissant, de l'envisager comme un outil de luttes et de création » (Adam, 2018, p. 50) afin qu'il puisse être utilisé comme outil de subversion. Or, comme le rappelle Lupton (2017), le temps, le corps et l'espace ne sont que peu pris en considération dans les recherches des littératies des données. C'est pourquoi elle a proposé le concept de « data sense » qui recentre les études autour du corps et incorpore les interactions des capteurs numériques (utilisés dans l'auto-suivi pour générer des données personnelles), les sens humains incarnés et la création de sens humain (Lupton *et al.*, 2018), permettant la production de « données vivantes » (*lively data*).
- 57 En reprenant ce concept, ces dernières seraient intégrées dans une vision « viscérale » des données (*data visceralizations*) où elles seraient « des représentations de l'information basées sur un certain nombre de sens comme le toucher, l'odorat ou le goût qui, en collaborant ensemble, stimulent nos sentiments et nos pensées²⁵ » (Stark, 2014).
- 58 Pour concevoir et récupérer ces données, Lettow (2011) a proposé le terme de « somatechnologie » qui désigne les technologies employant des « somatechniques » utilisées pour « mettre en évidence l'inextricabilité du *soma* et de la *technè*, du corps (en tant que construction culturellement intelligible) et des techniques (dispositifs et « technologies dures ») dans et à travers lesquelles les corps sont formés et transformés²⁶ ». C'est ainsi rappeler que la matérialité du corps est toujours médiée socialement et techniquement, loin d'une vision neutre des données.
- 59 Nous pouvons ici citer les travaux de Fraser (2021, 2024) sur les technologies immersives, notamment l'usage de la vidéo 360, qui permettent d'étudier les corporéités *queer* selon une démarche « ethnographique vidéographiée » située, offrant la possibilité de collecter des « petites données » qualitatives grâce à une captation des corps *queer* en mouvement, ce qui inclut également le corps du chercheur.
- 60 Enfin, la temporalité. Taylor (2010) souligne que les visions hétéronormatives font des vies une « course » réalisée de manière linéaire, séquentielle, à savoir : la naissance, l'enfance, l'adolescence, le début de la vie d'adulte, le mariage, la reproduction, l'éducation des enfants, la retraite, la vieillesse, la mort et l'héritage. Les vies *queer* à l'inverse se soustraient à la linéarité, bien qu'un phénomène normalisateur ait lieu (comme la volonté d'accéder au mariage et d'avoir des enfants), et s'incarne dans une autre temporalité (comme le passage du « coming-out » dans un contexte occidental). Plusieurs auteurs ont ainsi proposé de nouveaux cadres d'analyse, à l'instar de Freccero (2007) et sa « spectralité *queer* » qui se présente comme une temporalité non linéaire qui interroge les attachements mélancoliques de certains contre-récits *queer* et l'illusion d'un choix « vie/mort » à faire. Le tout au service d'une *queerisation* de l'historiographie offrant une relation plus éthique au passé et contrant la fiction symptomatique d'un avenir reproductif (*symptomatic fantasy of reproductive futurity*).

L'ouvrage d'Idier et Chauvin (2025) participe également à ce contre-récit en mettant en lumière une histoire « oubliée » de la persécution de l'homosexualité par le système policier et judiciaire français sur la base de rapports exhumés d'archives.

L'approche intersectionnelle

- 61 L'intersectionnalité, terme conçu par Crenshaw (2023) prend sa source dans les « études sur les femmes noires » (*Black Women's studies*) et dans « l'histoire intellectuelle des Noires » (p. 10-11) et vise « à nommer l'expérience des femmes noires et à améliorer leur qualité de leur vie comme de celle de toutes les personnes exclues, marginalisées ou punies par les lois qui auraient dû les protéger » (*ibid.*, p. 12), de faire valoir comment « race, genre et classe affectent le vécu des femmes noires aux États-Unis, en particulier sur le plan professionnel et légal » (*ibid.*, p. 10).
- 62 L'intersectionnalité pense la discrimination autrement que par un prisme factoriel singulier (*ibid.*, p. 13), car son expérience dépasse le racisme et le sexisme, Le concept s'est par la suite étendu à un ensemble varié de thématiques. Cette approche permettrait d'éviter les « silences construits » (De Lauretis, 1991) qui invisibilisent les corpus *queer*, notamment racisés (Back, 2024), et nécessiterait une collaboration constante de plusieurs disciplines et philosophies.
- 63 Comme évoqué, la littérature des données *queer* ne peut se suffire à elle seule. Elle doit permettre la rencontre croisée d'études sur les données liées au féminisme, au racisme, au handicap et toute autre situation exposant aux discriminations. Sur ce point, nous pouvons citer les travaux de McGlotten (2016) qui mêlent *Black studies* et *queer* dans la conception des données *queer*, reliant la notion de *Black data* aux concepts d'effacement, d'opacité et de cryptage.

L'approche éthique

- 64 La dernière approche est éthique. Basée sur les propositions de Guyan (2022) et Govender et Andrews (2021), la littérature des données *queer* doit intégrer une version spécifique de l'éthique des données, censée introduire et évaluer des principes intervenant sur tout le cycle de vie des données pour aider à comprendre leurs incidences sur la société (Arruda et Gerrits, 2020) et in fine, prendre des décisions « éclairées ». Le seul exemple de synthèse de la littérature est celui d'Harfoursh (2021) où les principes de transparence, du choix, du consentement, de la sécurité, de l'intégrité, de l'accès et de la responsabilité sont présentés. Principes auxquels Pothier et Condon (2020) ajoutent la promotion et la préservation d'un engagement responsable avec l'environnement ; ce dernier désignant le respect d'une part des personnes qui nous entourent et d'autre part des pratiques écologiques. Sur ce dernier point, nous spécifions celle des vies *queer* par l'adoption des points suscités dans la création et l'analyse des données, mais également celle des dites données, en particulier dans un contexte politique où celles-ci tendant à être invisibilisées, voire supprimées des bases de données comme dans le cas du « big delete » des politiques états-uniennes citées en introduction.
- 65 Guyan (2022) ajoutait par ailleurs la nécessité d'une vision éthique des finalités de projets de recueil de données, remettant ici en doute les visions « objectives » et « sincères » des méthodologies classiques²⁷, en listant un certain nombre de questions

centrales permettant de déterminer l'objectif et la transparence des études, leur restitution authentique et leur potentiel bénéfique pour les personnes *queer*.

- 66 Vis-à-vis des méthodologies scientifiques classiques évoquées ci-dessus, Detamore (2010) proposait enfin une approche éthique *queer* en tant que « méthode » explorant les politiques de l'intimité. Il préconisait le développement d'une relation d'intimité permettant la coproduction (1) des savoirs de manière horizontale qui légitime les visions et ressentis des enquêtés et lie ces derniers au chercheur, et (2) de nouveaux espaces politiques faisant émerger des mondes sociaux alternatifs.

Les limites d'une approche *queer*

- 67 Du fait de son « potentiel indélébile pour l'inversion/subversion » (Allen, 2011), de sa fluidité inhérente qui fait des sujets et des subjectivités des « fluides, instables et en perpétuel devenir », la « queerisation » des données reste encore sur ce point une interrogation. Duncan (2021) rappelait que le choix du terme *queer* n'est pas sans conséquences, étant nécessairement limité pour représenter les vies *queer*.
- 68 Parmi les limites citées dans la littérature se trouve la fluidité inhérente au concept de *queer* (Allen, 2011). Cette fluidité rend difficile le recueil de données et les significations que l'on peut en tirer (Nash, 2016) : en effet, la décomposition constante des catégories ne permet pas d'établir des données de référence, et par extension toute allocation de ressources pour des recherches, laissant un doute sur le sens que l'on peut tirer d'elles (Browne, 2010).
- 69 Par ailleurs, le *queer* pose une problématique d'historicité et d'archivage des données : si celles-ci changent constamment, quelle trace conserver ? Le cas s'est posé avec Wikipédia pour les fiches des personnes transgenres, en particulier sur le nom à indiquer. Actuellement, le *deadname* (le prénom de naissance) est indiqué dans l'infoboîte et le nom actuellement utilisé par la personne sert de titre à la page. La littératie des données distribuée citée plus tôt apporte une première réponse pour les deux points suscités en intégrant la réalisation systématique d'une documentation pour chaque production de données, apportant un élément de contexte pérenne aux transformations que lesdites données peuvent subir.
- 70 Le *queer* soulève également la question de la personne désignée pour recueillir et traiter les données. Pour Guyan (2022), même une personne auto-déclarée *queer* n'est pas un gage suffisant du fait que son identité et son orientation sont également jumelées à d'autres positions qui peuvent influencer ses pensées et objectifs (notamment le fait d'être une personne ou non cisgenre). Les travaux intégrant une approche située (Haraway, 2007) sont ici une proposition pour sortir d'une supposée neutralité axiologique, comme ceux de Detamore (2010) et Fraser (2021, 2024) présentés plus haut, et ainsi déterminer les limites précises des approches engagées pour mieux les compléter (ex. : intégration de profils plus en lien avec la thématique traitée dans les données).

Conclusion

- 71 Cette proposition de littératie des données *queer* s'adresse autant aux personnes *queer* que non *queer* dans le développement d'une prise de conscience des enjeux des données

et de leur portée sur les vies *queer*. S'appuyer sur les méthodes présentées ici doit nous permettre de tendre vers l'avènement de lettré.es des données *queer* ayant développé une culture technique leur permettant de traiter ces thématiques sans contraintes ni dépendances. C'est s'inscrire dans les études critiques en assumant un rôle engagé, politique, dans la conception et la préservation des données *queer*, à contre-courant de pratiques techniques essentialisantes issues de la *datafication* et de politiques toujours plus réactionnaires qui mettent en danger les vies *queer*. Une application de cette quintuple approche dans des projets liés à la conception et au traitement des données *queer* fera l'objet de futures publications.

BIBLIOGRAPHIE

- ADAM Zoé (2018), *Praxis Queer : Les corps queers comme sites de création et de résistance*, Thèse de doctorat sous la direction de Sam BOURCIER, Lille, Université de Lille, <https://theses.fr/2018LILUH034>, page consultée le 24 novembre 2025.
- ADAMS Tony, Carolyn ELLIS et Stacy Holman JONES (2017), « Autoethnography », *The International Encyclopedia of Communication Research Methods*, p. 1-11, DOI : <https://doi.org/10.1002/9781118901731.iecrm0011>.
- AHMED Sara (2012), *On Being Included: Racism and Diversity in Institutional Life*, Durham, Duke University Press Books.
- AKRICH Madeleine (1987), « Comment décrire les objets techniques ? », *Techniques et Culture*, 9, p. 49-64.
- ALESSANDRIN Arnaud et Anastasia MEIDANI (2022), « Santé des personnes non binaires : de quoi parle-t-on ? », *Santé Publique*, 34(HS2), p. 97-102, DOI : <https://doi.org/10.3917/spub.hs2.0097>.
- ALLEN Louisa (2011), « Bending the yules: Attempting Queer research on sexuality in schools », *Gay & Lesbian Issues and Psychology Review*, 7(2), p. 155-167.
- ANDERSON Gary et IRVINE Patricia (1993), « Informing Critical Literacy with Ethnography », dans *Critical Literacy: Politics, Praxis, and the Postmodern*, États-Unis, State University of New York, p. 81-104.
- ARRUDA Sergie et GERRITS Andrei (2020), « Exploring data literacy: Concepts and determinants for data skills development », Dissertation, Coimbra, Université de Coimbra, <https://estudogeral.sib.uc.pt/handle/10316/94716>, page consultée le 02 décembre 2025.
- BACHIMONT Bruno (2020), « Rhétorique de l'intelligibilité, vers des humanités du numérique - NumeRev », *NumeRev*, 28 février, <https://projet.numerev.com/lab/65-rhetorique-de-l-intelligibilite-vers-des-humanites-du-numerique>, page consultée le 30 novembre 2025.
- BACK Camille (2024), « Nos “silences construits” : nommer la violence épistémique au fondement de la théorie queer », *Genre, sexualité & société*, 31, DOI : <https://doi.org/10.4000/122tb>.
- BAKHTINE Mikhaïl, Michael HOLQUIST et Caryl EMERSON (1982), *The Dialogic Imagination: Four Essays*, Austin, University of Texas Press.

- BALANDIER Georges (2001), *Le grand système*, Paris, Fayard.
- BELCASTRO Celeste (1994), *Detecting Upset in Fault Tolerant Control Computers Using Data Fusion Techniques*, Drexel, Université de Drexel, DOI : <https://doi.org/10.17918/00000473>.
- BERNINI Lorenzo (2024), *Philosophies Queer*, Les Lilas, Eterotopia.
- BORDAS Éric (2024), « Cultures gays et identités queers, ou la pensée d'un style », *Critique* 683(4), p. 280-295, DOI : <https://doi.org/10.3917/criti.683.0280>.
- BORGHI Rachele (2012), « De l'espace genré à l'espace queerisé. Quelques réflexions sur le concept de performance et son usage en géographie », *Travaux et documents de ESO* 33, p. 109-116.
- BORGHI Rachele (2014), « L'espace à l'époque du queer : contaminations queer dans la géographie française », *Politiqueer*, Dimensions francofolles, <https://hal.science/hal-01382341>, page consultée le 2 décembre 2025.
- BOUQENTAR Lamiae (2021), « Écriture auto-ethnographique et positionnement : réflexivité entourant l'arabité et le queer », *Alterstice : revue internationale de la recherche interculturelle*, 10(1), p. 95-107, DOI : <https://doi.org/10.7202/1084805ar>.
- BOURCIER Sam (2018), *Queer zones redux. La trilogie*, Paris, Amsterdam.
- BRAND Jess et Ina SANDER (2020), *Critical Data Literacy Tools for Advancing Data Justice: A Guidebook*, Data justice lab, <https://datajusticelab.org/wp-content/uploads/2020/06/djl-data-literacy-guidebook.pdf>, page consultée le 2 décembre 2025.
- BRIDGES Lauren (2021), « Digital failure: Unbecoming the "good" data subject through entropic, fugitive, and queer data », *Big Data & Society*, 8(1), DOI : <https://doi.org/10.1177/2053951720977882>.
- BRITZMAN Deborah (1995), « Is there a queer pedagogy? Or, stop reading straight », *Educational Theory*, 45(2), p. 151-165, DOI : <https://doi.org/10.1111/j.1741-5446.1995.00151.x>.
- BROWNE Kath (2010), « Queer quantification or queer(y)ing quantification: creating lesbian, gay, bisexual or heterosexual citizens through governmental social research », dans *Queer Methods and Methodologies: Intersecting Queer Theories and Social Science Research*, Londres, Routledge, p. 231-250, DOI : <https://doi.org/10.4324/9781315603223-15>.
- BRYSON Mary et Suzanne DE CASTELL (1993), « Queer pedagogy: Praxis makes im/perfect », *Canadian Journal of Education/Revue Canadienne de l'éducation*, 18(3), p. 285-305.
- BUMAYE Dawud (2024), « Construire des théories queer of color en France », *Genre, sexualité & société*, 3, DOI : <https://doi.org/10.4000/122tf>.
- BUTLER Judith (2011), *Bodies That Matter: On the Discursive Limits of Sex*, 1ère édition, Londres, Routledge.
- BUTLER Judith (2006), *Trouble dans le genre : Le féminisme et la subversion de l'identité*, Paris, La Découverte.
- BUTLER Judith (1997), « Critically Queer », dans Shane PHELAN, *Playing With Fire*. Londres, Routledge.
- CARDON Dominique (2012), « Regarder les données », *Multitudes*, 49(2), p. 138-142, DOI : <https://doi.org/10.3917/mult.049.0138>.
- CARLSON Jake, Michael FOSMIRE, Chris MILLER et Megan SAPP NELSON (2011), « Determining data information literacy needs: A study of students and research faculty », *Libraries Faculty and*

Staff Scholarship and Research, West Lafayette, Purdue University, https://docs.lib.purdue.edu/lib_fsdocs/23, page consultée le 2 décembre 2025.

CASEMAJOR Nathalie (2015), « Digital materialisms: Frameworks for digital media studies », *Westminster Papers in Communication and Culture*, 10(1), DOI : <https://doi.org/10.16997/wpcc.209>.

CASTELLS Manuel (2009), *The Power of Identity*. 2eme édition, Malden, Wiley-Blackwell.

CERVULLE Maxime (2023), « Le féminisme comme théorie de la représentation : Teresa de Lauretis de l'autre côté du miroir », *Le genre du monde*, p. 173-197.

CERVULLE Maxime et Franck FREITAS (2015), *Poli N°10 : Techno-racismes*, Paris, Poli Politique.

CERVULLE Maxime et Nelly QUEMENER (2021), « Queer », dans Juliette RENNES (dir.), *Encyclopédie critique du genre*, Paris, La Découverte, p. 632-642.

CITTON Yves (2017), « L'initiation numérique : de l'entreprise des données à la grâce du commun », dans Isabelle MOREL et Joël MOLINARO (dir.), *Être initié à l'heure des mutations anthropologiques*, Paris, Les éditions du Cerf, p. 207-226.

COULDRY Nick et Ulises MEJIAS (2019), « Data colonialism: rethinking big data's relation to the contemporary subject », *Television & New Media*, 20(4), p. 336-349, DOI : <https://doi.org/10.1177/1527476418796632>.

CRAWFORD Kate (2022), *Contre-atlas de l'intelligence artificielle : Une cartographie politique, sociale et environnementale de l'IA*, traduit par Laurent BURY, Paris, Zulma.

CRAWFORD Kate (2013), « The hidden biases in big data », *Harvard Business Review*, 1er avril, <https://hbr.org/2013/04/the-hidden-biases-in-big-data>.

CREMONINI Marco (2023), *A Critical Take on Privacy in a Datafied Society*, DOI : <https://doi.org/10.48550/arXiv.2308.02573>.

CRENSHAW Kimberlé et Emmanuelle DELANOE (2023), *Intersectionnalité : Deux essais*, Paris, Payot.

DAMIAN-GAILLARD Béatrice et Florian VOROS (2023), « Du discours au dispositif : Penser la fabrique médiatique des sexualités », *Réseaux*, 237(1), p. 9-34, DOI : <https://doi.org/10.3917/res.237.0009>.

DE LAURETIS Teresa (1987), *Technologies of Gender: Essays on Theory, Film, and Fiction*, Bloomington, Indiana University Press.

DE LAURETIS Teresa (1991), « Queer theory: Lesbian and gay sexualities an introduction », *differences*, 3(2), p. iii-xviii, DOI : <https://doi.org/10.1215/10407391-3-2-iii>.

DE LAURETIS Teresa, Pascale MOLINIER, Maxime CERVULLE et Sam BOURCIER (2023), *Théorie queer et cultures populaires : De Foucault à Cronenberg*, Paris, La Dispute.

DELPHY Christine (2001), *L'Ennemi principal, tome 2 : Penser le genre*, Paris, Editions Syllepse.

DENCIK Lina, Arne HINTZ, Joanna REDDEN et Emiliano TRERE (2022), *Data Justice. 1ère édition*, London Thousand Oaks, Sage Publications Ltd.

DENCIK Lina, Arne HINTZ, Joanna REDDEN et Emiliano TRERE (2019), « Exploring data justice: Conceptions, applications and directions », *Information, Communication & Society*, 22(7), p. 873-881, DOI : <https://doi.org/10.1080/1369118X.2019.1606268>.

DETAMORE Mathias (2010), « Queer(y)ing the ethics of research methods: Toward a politics of intimacy in researcher/researched relations », dans *Queer Methods and Methodologies*, Londres, Routledge.

- D'IGNAZIO Catherine (2024), *Counting Feminicide: Data Feminism in Action*, The MIT Press.
- D'IGNAZIO Catherine (2017), « Creative data literacy: Bridging the gap between the data-haves and data-have nots », *Information Design Journal*, 23(1), p. 6-18, DOI : <https://doi.org/10.1075/idj.23.1.03dig>.
- D'IGNAZIO Catherine et Lauren KLEIN (2020), *Data Feminism*, Cambridge, Massachusetts, The MIT Press.
- DORLIN Elsa (2008), *Sexe, genre et sexualités*, Paris, Presses universitaires de France.
- DRUCKER Johanna (2011), « Humanities Approaches to Graphical Display », *Digital Humanities Quarterly*, 5(1), p. 1-21.
- DUNCAN Laura (2021), « Queer data: Medical quantification and what counts about counting », thèse de doctorat sous la direction de Kelly KNIGHT, San Francisco, Université de Californie, <https://www.proquest.com/docview/2555593344/abstract/6314572AA4F648CDPQ/1>, page consultée le 2 décembre 2025.
- DYMYTROVA Valentyna (2011), « Identités politiques, discours et médias : Le cas de la "Révolution Orange" », thèse de doctorat sous la direction de Bernard LAMIZET, Lyon, Université Lyon 2, https://theses.hal.science/tel-01376987/file/these_internet_dymytrova_v_vol1.pdf, page consultée le 2 décembre 2025.
- EDWARD Mark et Chris GREENOUGH (2019), « Queer Literacy », dans *Handbook of Research Ethics and Scientific Integrity*, édité par Ron IPHOFEN, Cham, Springer International Publishing, p. 1-12, DOI : https://doi.org/10.1007/978-3-319-76040-7_57-1.
- ELLUL Jacques et Jean-Luc PORQUET (2012), *Le système technicien*, Paris, Cherche midi.
- ELLUL Jacques (1988), *Le Bluff technologique*, Paris, Hachette.
- ERIBON Didier (2003), *Dictionnaire des cultures gays et lesbiennes*, Paris, Larousse.
- FLIPO Fabrice (2020), *L'impératif de la sobriété numérique : L'enjeu des modes de vie*, Paris, Editions Matériologiques.
- FLORIDI Luciano (2009), « Philosophical conceptions of information », dans Giovanni SOMMARUGA (dir.), *Formal Theories of Information: From Shannon to Semantic Information Theory and General Concepts of Information*, Berlin, Heidelberg : Springer, p. 13-53, DOI : https://doi.org/10.1007/978-3-642-00659-3_2.
- FOTOPOULOU Aristeia (2020), « Conceptualising critical data literacies for civil society organisations: Agency, care, and social responsibility », *Information, Communication & Society*, p. 1-18, DOI : [10.1080/1369118X.2020.1716041](https://doi.org/10.1080/1369118X.2020.1716041).
- FOUCAULT Michel (1984), « Espace, savoir et pouvoir », dans *Dits Écrits tome IV*, Paris, Gallimard.
- FOUCAULT Michel (1994), *Histoire de la sexualité, I : La volonté de savoir*, Paris, Gallimard.
- FRASER Jordan Emery (2024), « Corps, espace et numérique : les potentialités queer contemporaines dans la constitution de contre-espaces politiques. Études ethno-graphiques 360° au Brésil et en France », Thèse de doctorat sous la direction de Jacques IBANEZ BUENO et Arley ANDRIOLO, Chambéry, Université Savoie Mont Blanc, <https://theses.fr/2024CHAMA052>, page consultée le 3 décembre 2025.
- FRASER Jordan Emery (2021), « Dispositif immersif en contexte queer », *Revue française des méthodes visuelles*, 5, DOI : <https://doi.org/10.4000/12mpp>.

FRECCERO Carla (2007), « Queer spectrality: Haunting the past », dans *A Companion to Lesbian, Gay, Bisexual, Transgender, and Queer Studies*, New York, John Wiley & Sons, p. 194-213, DOI : <https://doi.org/10.1002/9780470690864.ch10>.

GITELMAN Lisa (2013), « *Raw Data* » *Is an Oxymoron*, Cambridge, The MIT Press.

GOSSETT Reina, Eric STANLEY et Johanna BURTON (2017), *Trap door: Trans cultural production and the politics of visibility*, Cambridge, The MIT Press.

GOVENDER Navan et Grant ANDREWS (2021), « Queer critical literacies », dans Jessica ZACHER PANDYA, Raúl Alberto MORA, Jennifer Helen ALFORD, Noah Asher GOLDEN et Roberto Santiago DE ROOCK (dir.), *The Handbook of Critical Literacies*, Londres, Routledge.

GUYAN Kevin (2022), *Queer Data*, Londres, Bloomsbury Publishing PLC.

HABERMAS Jurgen (1988), *L'Espace public*, Paris, Payot.

HARAWAY Donna, Laurence ALLARD, Delphine GARDEY et Nathalie MAGNAN (2007), *Manifeste cyborg et autres essais : Sciences - Fictions - Féminismes*, Paris, Éditions Exils.

HARFOUSH Fady (2021), « Data ethics: Facts and fiction », dans Fady HARFOUSH (dir.), *From Big Data to Intelligent Data: An Applied Perspective*, Cham, Springer International Publishing, p. 87-99, DOI : https://doi.org/10.1007/978-3-030-76990-1_7.

HAUTEA Samantha, Sayamindu DASGUPTA et Benjamin MAKO HILL (2017), « Youth perspectives on critical data literacies », dans *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI'17, New York, Association for Computing Machinery, p. 919-930, DOI : <https://doi.org/10.1145/3025453.3025823>.

HELMER Kirsten (2016), « Reading queer counter-narratives in the high-school literature classroom: possibilities and challenges », *Discourse: Studies in the Cultural Politics of Education*, 37(6), p. 902-916, DOI : <https://doi.org/10.1080/01596306.2015.1120943>.

HEPP Andreas, Juliane JARKE et Leif KRAMP (2022), « New perspectives in critical data studies: The ambivalences of data power — An introduction », dans Andreas HEPP, Juliane JARKE et Leif KRAMP (dir.), *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, Transforming Communications — Studies in Cross-Media Research. Cham, Springer International Publishing, p. 1-23, DOI : https://doi.org/10.1007/978-3-030-96180-0_1.

HJØRLAND Birger (2018), « Data (with big data and database semantics) », *Knowledge Organization*, 45(8), p. 685-708, DOI : <https://doi.org/10.5771/0943-7444-2018-8-685>.

HOFFMANN Anna Lauren (2021), « Terms of inclusion: Data, discourse, violence », *New Media & Society*, 23(12), p. 3539-3556, DOI : <https://doi.org/10.1177/1461444820958725>.

IDIER Antoine et Sébastien CHAUVIN (2025), *Réprimer et réparer ; une histoire effacée de l'homosexualité*, Paris, Textuel.

IJJASZ-VASQUEZ Ede et Clifton CORTEZ (2017), « To fight discrimination, we need to fill the LGBTI data gap », *World Bank Blogs* (blog), <https://blogs.worldbank.org/en/voices/fight-discrimination-we-need-fill-lgbti-data-gap>, page consultée le 2 décembre 2025.

IRGENS Golnaz Arastoopour, Simon KNIGHT, Alyssa WISE, Thomas PHILIP, Maria OLIVARES, Sarah VAN WART...Jennifer KAHN (2020), « Data literacies and social justice: Exploring critical data literacies through sociocultural perspectives », *Learning and Identity*, juin, <https://repository.isls.org/handle/1/6665>, page consultée le 2 décembre 2025.

JAHHAH Marc (2024), « Écrire et intervenir en scientifique queer de couleur : une épopée des réseaux », *Genre, sexualité & société*, 31, DOI : <https://doi.org/10.4000/122tc>.

- JAHJAH Marc (2022), « “T’es intelligent pour un arabe !” Auto-ethnographie d’un corps colonisé », *Itinéraires. Littérature, textes, cultures*, 2021(3), DOI : <https://doi.org/10.4000/itineraires.11748>.
- JANICAUD Dominique (1985), *La Puissance du rationnel*, Paris, Gallimard.
- JANKS Hilary (2013), « Critical literacy in teaching and research 1 », *Education Inquiry*, 4(2), p. 225-242, DOI : <https://doi.org/10.3402/edui.v4i2.22071>.
- JONES Nicola (2024), « The AI revolution is running out of data. what can researchers do? », *Nature*, 636(8042), p. 290-292, DOI : <https://doi.org/10.1038/d41586-024-03990-2>.
- JUNG Gustav (1996), *Problèmes de l’âme moderne*, Paris, Buchet-Chastel.
- JUREK Lucie, Ludovic SOUILLER et François MEDJKANE (2022), « L’identité de genre, entre faits naturels et faits construits, une approche intégrative et développementale », *médecine/sciences*, 38(10), p. 808-815, DOI : <https://doi.org/10.1051/medsci/2022128>.
- KEYES Os (2019), « Counting the Countless », *Real Life*, <https://reallifemag.com/counting-the-countless/>, page consultée le 2 décembre 2025.
- KEYES Os (2018), « The misgendering machines: Trans/hci implications of automatic gender recognition », *Proc. ACM Hum.-Comput. Interact*, 2(CSCW), <https://doi.org/10.1145/3274357>.
- KIRSCHENBAUM Matthew, Richard OVENDEN et Gabriela REDWINE (2010), *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*, Washington, Council on Library and Information Resources Washington.
- KITCHIN Rob (2014), *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*, Los Angeles, SAGE Publications Ltd.
- KOKAS Aynne (2023), « Data trafficking and the international risks of surveillance capitalism: The case of grindr and china », *Television & New Media*, 24(6), p. 673-690, DOI : <https://doi.org/10.1177/15274764221137250>.
- LAMIZET Bernard (2002), *Politique et identité*, Lyon, Presses Universitaires de Lyon.
- LANDSTROM Catharina (2007), « Queering Feminist Technology Studies », *Feminist Theory*, 8(1), p. 7-26, DOI : <https://doi.org/10.1177/1464700107074193>.
- LANIGAN Richard, « Capta versus Data: Method and evidence in communicology », *Human Studies*, 17(1), p. 109-130.
- LATOUR Bruno (2007), « Pensée retenue, pensée distribuée », dans Christian JACOB (dir.), *Lieux de savoir 1. Espaces et communautés*, Paris, Albin Michel.
- LE DEUFF Olivier et Rayya ROUMANOS (2022), « Enjeux définitionnels et scientifiques de la littérature algorithmique : entre mécanologie et rétro-ingénierie documentaire », *tic&société*, 15(2-3), p. 325-360, DOI : <https://doi.org/10.4000/ticetsociete.7105>.
- LE DEUFF Olivier (2026), « Des données à l’objet. Les POCL (Petits objets connectés ludiques) entre culture technique et œuvre ouverte », *Vingt-quatrième conférence internationale sur le document électronique (CIDE.24), actes du 20 et 21 novembre 2025*, Bordeaux, Édition Europia.
- LELEU-MERVIEL Sylvie (2016), « La signifiante canalisée par l’horizon de pertinence, des saisies aux agrégats via les données », *Revue des Interactions Humaines Médiatisées (RIHM) = Journal of Human Mediated Interactions*, 17, p. 107-139, <https://hal.science/hal-03184831/>, page consultée le 3 décembre 2025.
- LETTOW Susanne (2011), « Somatechnologies: Rethinking the body in the philosophy of technology », *Techné*, 15, p. 110-17.

LEURS Koen et Tamara SHEPERD (2017), « 15. Datafication & discrimination », dans Mirko Tobias SCHAFER et Karin VAN ES (dir.), *The Datafied Society: Studying Culture Through Data*, Amsterdam, Amsterdam University Press, p. 211-232, DOI : <https://doi.org/10.1515/9789048531011-018>.

LUHMANN Susanne (1998), « Queering/Querying pedagogy? Or, pedagogy is a pretty queer thing », dans *Queer Theory in Education*, Londres, Routledge, p. 141-155.

LUPTON Deborah (2017), « Feeling your data: Touch and making sense of personal digital data », *New Media & Society* 19(10), p. 1599-1614, DOI : <https://doi.org/10.1177/1461444817717515>.

LUPTON Deborah, Sarah PINK, Christine HEYES LABOND et Shanti SUMARTOJO (2018), « Digital traces in context| personal data contexts, data sense and self-tracking cycling », *International Journal of Communication*, 12, p. 19.

MARIGNIER Noémie (2016), « Les matérialités discursives du sexe : la construction et la déstabilisation des évidences du genre dans les discours sur les sexes atypiques », Thèse de doctorat sous la direction de Marie-Anne PAVEAU et Luca GRECO, Paris, Sorbonne Paris Cité, <https://theses.fr/2016USPCD057>, page consultée le 2 décembre 2025.

MAYER-SCHONBERGER Viktor et Kenneth CUKIER (2013), *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Boston, Eamon Dolan / Houghton Mifflin Harcourt.

MCGLOTTEN Shaka (2016), « Black Data », dans Patrick JOHNSON (dir.), *No Tea No Shade*, Duke University Press.

MEIDANI Anastasia et Arnaud ALESSANDRIN (2019), « Cancer, territoires et discrimination : des zones rurales aux quartiers prioritaires de la ville », *Santé Publique*, 31(5), p. 693-702, DOI : [tps://doi.org/10.3917/spub.195.0693](https://doi.org/10.3917/spub.195.0693).

MELLINI Laura (2009), « Entre normalisation et hétéronormativité : la construction de l'identité homosexuelle », *Déviance et Société*, 33(1), p. 3-26, DOI : <https://doi.org/10.3917/ds.331.0003>.

MIELI Mario, Gianni ROSSI BARILLI et Paola MIELI (2017), *Elementi di critica omosessuale*, Saggi, Universale Economica.

MILAN Stefania et Lonneke van der VELDEN (2016), « The Alternative Epistemologies of Data Activism », *Digital Culture & Society*, 2(2), p. 57-74, DOI : <https://doi.org/10.14361/dcs-2016-0205>.

MILIONI Dimitra et Venetia PAPA (2022), « The oppositional affordances of data activism », *Media International Australia*, 183(1), p. 44-59, DOI : <https://doi.org/10.1177/1329878X221074795>.

MILLER Sj (2015), « A Queer literacy framework promoting (a)gender and (a)sexuality self-determination and justice », *English Journal*, 104(5), p. 37-44, DOI : <https://doi.org/10.58680/ej201527209>.

MONEY John (1955), « Hermaphroditism, Gender and Precocity in Hyperadrenocorticism: Psychologic Findings », *Bulletin of the Johns Hopkins Hospital*, 96(6), p. 253-264.

MONRO Surya (2019), « Non-binary and genderqueer: An overview of the field », *International Journal of Transgenderism*, 20, p. 1-6, DOI : <https://doi.org/10.1080/15532739.2018.1538841>.

MOSCO Vincent (2014), *To the Cloud: Big Data in a Turbulent World*, Londres, Routledge.

NASH Catherine (2016), *Queer Methods and Methodologies: Intersecting Queer Theories and Social Science Research*, Londres, Routledge. DOI : <https://doi.org/10.4324/9781315603223>.

NEMI NETO João (2018), « Queer pedagogy: Approaches to inclusive teaching », *Policy Futures in Education*, 16(5), p. 589-604, DOI : <https://doi.org/10.1177/1478210317751273>.

NIEDERGANG Pierre (2023), *Vers la normativité queer*, Toulouse, Blast.

- NOBLE Safiya Umoja (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, NYU Press.
- NORRIS Stephen et Linda PHILLIPS (2003), « How literacy in its fundamental sense is central to scientific literacy », *Science Education*, 87(2), p. 224-240, DOI : <https://doi.org/10.1002/sce.10066>.
- OLLIVIER Bruno (2007), *Identité et identification : sens, mots et techniques*, Paris, HERMES.
- OLSON David (1998), *L'univers de l'écrit ; comment la culture écrite donne forme à la pensée*, Paris, Editions Retz, <https://www.boitealivres.com/livre/9782725629926-l-univers-de-l-ecrit-comment-la-culture-ecrite-donne-forme-a-la-pensee-david-r-olson/>, page consultée le 3 décembre 2025.
- OLSON David et Jean-Claude LEJOSNE (2006), « Littératie, scolarisation et cognition. Quelques implications de l'anthropologie de Jack Goody », *Pratiques*, 131(1), p. 83-94, DOI : <https://doi.org/10.3406/prati.2006.2119>.
- O'NEIL Cathy (2022), *The Shame Machine: Who Profits in the New Age of Humiliation*, New York, Crown.
- O'NEIL Cathy (2017), *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York, Crown.
- ORTIZ-FREULER Juan (2023), « Datafication, identity, and the reorganization of the category individual », Rochester, Social Science Research Network, <https://papers.ssrn.com/abstract=4501828>, page consultée le 2 décembre 2025.
- OUDIN Fanny (2012), « Identité et persona. Quelques réflexions liminaires autour de l'image de soi au Moyen Âge », *Questes. Revue pluridisciplinaire d'études médiévales*, 24, p. 27-47, DOI : <https://doi.org/10.4000/questes.3090>.
- PERREAU Bruno (2018), *Qui a peur de la théorie queer ?*, Paris, Les Presses de Sciences Po.
- PHILIP Thomas, Maria OLIVARES-PASILLAS et Janet ROCHA (2016), « Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations », *Cognition and Instruction*, 34(4), p. 361-388, DOI : <https://doi.org/10.1080/07370008.2016.1210418>.
- POIRIER-SAUMURE Alexis (2024), « Qu'y a-t-il de queer dans la pédagogie queer actuelle ? », *Arborescences : revue d'études littéraires, linguistiques et pédagogiques de langue française*, 14, p. 23-36, DOI : <https://doi.org/10.7202/1113547ar>.
- POTIER Wendy Girven et Patricia CONDON (2020), « Towards data literacy competencies: Business students, workforce needs, and the role of the librarian », *Journal of Business & Finance Librarianship*, 25(3-4), p. 123-146, DOI : <https://doi.org/10.1080/08963568.2019.1680189>.
- POTTER Troy (2022), « Queer literacies: Queer objects, disorientation, and sponsors of literacy », *Journal of Literacy Research*, 54(1), p. 98-106, DOI : <https://doi.org/10.1177/1086296X221076452>.
- PRECIADO Beatriz (2003), « Multitudes queer : Notes pour une politiques des "anormaux" », *Multitudes*, 12(2), p. 17-25, DOI : <https://doi.org/10.3917/mult.012.0017>.
- PRIVAT Jean-Marie (2010), « Un bain de littératie. A l'école de la piscine », *Ethnographiques.org : revue en ligne des sciences humaines et sociales*, 20, <https://www.ethnographiques.org/2010/Privat>. Page consultée le 3 décembre 2025.
- PROULX Serge (2020), *La participation numérique : une injonction paradoxale*, Paris, Presse des Mines, <https://www.pressesdesmines.com/produit/la-participation-numerique-une-injonction-paradoxale/>.

QUIJANO Aníbal (2007), « “Race” et colonialité du pouvoir », *Mouvements*, 51(3), p. 111-118, DOI : <https://doi.org/10.3917/mouv.051.0111>.

REYNOLDS Paul (2001), « Accounting for sexuality: The scope and limitations of census data on sexual identity and difference », *Radical Statistics*, 78, <https://www.radstats.org.uk/no078/reynolds.htm>, page consultée le 3 décembre 2025.

RICAUORTE Paola (2019), « Data epistemologies, the coloniality of power, and resistance », *Television & New Media*, 20(4), p. 350-365. DOI : <https://doi.org/10.1177/1527476419831640>.

RICHTERICH Annika (2018), *The Big Data Agenda: Data Ethics and Critical Data Studies*. Westminster, University of Westminster Press, <https://library.oapen.org/handle/20.500.12657/30155>, page consultée le 2 décembre 2025..

ROBERT Pascal (2020), *L'impensé numérique. Tome 2 : Interprétations critiques et logiques pragmatiques de l'impensé*, Paris, Éditions des archives contemporaines.

ROSENBERG Tiina (2008), « Locally Queer. A Note on the feminist genealogy of Queer Theory », *The Graduate Journal of Social Science, Special Issue on Queer Studies*, 5(2), p. 5-18.

RUBERG Bonnie et Spencer RUELOS (2020), « Data for queer lives : How LGBTQ gender and sexuality identities challenge norms of demographics », *Big Data & Society*, 7(1), DOI : <https://doi.org/10.1177/2053951720933286>.

SANDRETTO Susan (2018), « A case for critical literacy with queer intent », *Journal of LGBT Youth*, 15(3), p. 197-211, DOI : <https://doi.org/10.1080/19361653.2018.1466753>.

SCANDURRA Cristiano, Roberto VITELLI, Nelson Mauro MALDONATO, Paolo VALERO et Vincenzo BOCHICCHIO (2019), « Une étude qualitative sur le stress minoritaire ressenti par les personnes transgenres et de genre non conforme en Italie », *Sexologies*, 28(3), p. 142-146, DOI : <https://doi.org/10.1016/j.sexol.2019.05.001>.

SCHRAM Brian (2019), « Accidental orientations: Rethinking Queerness in archival times », *Surveillance & Society*, 17(5), p. 602-617, DOI : <https://doi.org/10.24908/ss.v17i5.8688>.

SCOTT Joan et Eleni VARIKAS (1988), « Genre : Une catégorie utile d'analyse historique », *Les cahiers du GRIF*, 37(1), p. 125-153, DOI : <https://doi.org/10.3406/grif.1988.1759>.

SEDGWICK Eve (1998), « Construire des significations queer », dans Didier ERIBON (dir.), *Les études gay et lesbiennes, Actes du colloque des 23 et 27 juin 1997*, Paris, Éditions du Centre Georges Pompidou, p. 109-116.

SIMEONI Fe, María MENENDEZ-BLANCO, Rupesh VYAS et Antonella DE ANGELI (2024), « Querying the quantification of the queer: Data-driven visualisations of the gender spectrum », dans *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, DIS '24, New York, Association for Computing Machinery, p. 3243-3256, DOI : <https://doi.org/10.1145/3643834.3660695>.

SIMONDON Gilbert (2024), *Du mode d'existence des objets techniques*, Paris, Champs essais.

SLEDD Andrew (1988), « Readin' not riotin': The politics of literacy (essay) », *College English*, 50(5), p. 495-508.

STARK Luke (2014), « Come on feel the data (and smell it) », *The Atlantic* (blog), <https://www.theatlantic.com/technology/archive/2014/05/data-visceralization/370899/>, page consultée le 2 décembre 2025.

STIEGLER Bernard (2008), « Psychopouvoir et guerre métapsychologique : la question du pharmakon », dans *Enfants turbulents : l'enfer est-il pavé de bonnes préventions ?*, Toulouse, Érès, p. 119-130.

STORNAIULO Amy (2020), « Authoring data stories in a media makerspace: Adolescents developing critical data literacies », *Journal of the Learning Sciences*, 29(1), p. 81-103, DOI : <https://doi.org/10.1080/10508406.2019.1689365>.

STREET Brian (2006), « Autonomous and ideological models of literacy : Approaches from new literacy studies », *Current Issues in Comparative Education*, 5(2), DOI : <http://doi.org/10.52214/jcice.v5i2.11369>.

SYKES Heather (2011), « Hetero- and homo-normativity: Critical literacy, citizenship education and queer theory », *Curriculum Inquiry*, 41(4), p. 419-432, DOI : <https://doi.org/10.1111/j.1467-873X.2011.00561.x>.

TACHEVA Zhasmina et Srividya RAMASUBRAMANIAN (2023), « Challenging AI empire: Toward a decolonial and queer framework of data resurgence », *Advance, mars*, DOI : <https://doi.org/10.31124/advance.22012724.v1>.

TAYLOR Jodie (2010), « Queer temporalities and the significance of “music scene” participation in the social identities of middle-aged queers », *Sociology*, 44(5), p. 893-907.

TAYLOR Linnet, Siddharth Peter DE SOUZA, Aaron MARTIN et Joan LOPEZ SOLANO (2025), « Governing artificial intelligence means governing data: (re)setting the agenda for data justice », *Dialogues on Digital Society*, DOI : <https://doi.org/10.1177/29768640241306800>.

THEVIOT Anaïs (2023), « Saisir les données en contexte », dans Anaïs THEVIOT, *Gouverner par les données ? Pour une sociologie politique du numérique*, Gouvernement en question(s), Lyon, ENS Éditions, p. 13-32, <http://books.openedition.org/enseditions/44838>, page consultée le 3 décembre 2025.

THRELKELD Aubry (2014), « A critical queer literacy approach to teaching children’s literature about same-sex parenting », dans Salika LAWRENCE, *Critical Practice in P-12 Education: Transformative Teaching and Learning*, IGI Global, p. 223-242, DOI : <https://doi.org/10.4018/978-1-4666-5059-6.ch010>.

TOMASEV Nenad, Kevin MCKEE, Jackie KAY et Shakir MOHAMED (2021), « Fairness for unobserved characteristics: Insights from technological impacts on queer communities », dans *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, p. 254-265, DOI : <https://doi.org/10.1145/3461702.3462540>.

TRIPALDI Laura et Muriel MORELLI (2024), *Gender Tech : Comment la technologie contrôle le corps des femmes*, Paris, Lux Editeur.

VAN HORN Selena (2022), « LGBTQ+ and queer literacy research: A review of pk-12 classrooms and teacher preparation », dans Selena VAN HORN (dir.), *Towards Queer Literacy in Elementary Education : Always Becoming Allies*, Cham, Springer International Publishing, p. 29-66, DOI : https://doi.org/10.1007/978-3-031-17087-4_3.

VERDI Ugo et Olivier LE DEUFF (2020), « La data literacy distribuée. Périmètres définitionnels, origines documentaires, perspectives réticulaires », *Les Cahiers du numérique*, 16(2-3-4), p. 137-173, <https://www.jle.com/10.3166/LCN.2020.006>, page consultée le 2 décembre 2025.

VERDI Ugo (2024), « À la recherche d’une maîtrise des données : une étude des discours scientifiques et professionnels sur la data literacy », Thèse de doctorat sous la direction d'Olivier LE DEUFF, Bordeaux, Université Bordeaux-Montaigne, <https://theses.fr/s257319>, page consultée le 2 décembre 2025.

VERDI Ugo (2023), « Quelle(s) réponse(s) à l'enjeu d'acculturation aux données ? Un état de l'art des caractéristiques de la data literacy », *Revue française des sciences de l'information et de la communication*, 26, DOI : <https://doi.org/10.4000/rfsic.14589>.

VERLAET Lise (2024), « Le numérique comme artefact technoscientifique : une “r-évolution” paradigmatique ? », *Revue Intelligibilité du numérique*, 6, <https://intelligibilite-numerique.numerev.com/numeros/n-6-2024/2931-le-numerique-comme-artefact-technoscientifique-une-r-evolution-paradigmatique>, page consultée le 2 décembre 2025.

VITALI-ROSATI Marcello (2024), *Éloge du bug : Être libre à l'époque du numérique*, Paris, Zones.

WANG Suisui (2024), « Situated Indications: Queer STS Experiments on Global Datafication », *Queer Data Studies*, https://www.academia.edu/122778910/Situated_Indications_Queer_STS_Experiments_on_Global_Datafication, page consultée le 2 décembre 2025.

WANG Yilun et Michal KOSINSKI (2018), « Deep neural networks are more accurate than humans at detecting sexual orientation from facial images », *Journal of Personality and Social Psychology*, 114(2), p. 246-257, DOI : <https://doi.org/10.1037/pspa0000098>.

WARGO Jon (2022), « Queer Literacies », dans Stephanie Anne SHELTON et Kamden K. STRUNK, *Encyclopedia of Queer Studies in Education*, p. 530-p. 334, DOI : https://doi.org/10.1163/9789004506725_104.

WILSON Ara (2019), « Queer Anthropology », *The Open Encyclopedia of Anthropology*, 31, <https://www.anthroencyclopedia.com/entry/queer-anthropology>, page consultée le 3 décembre 2025.

WOLFF Alex (2022), « Queer sousveillance : Publics, politics, and social media in South Korea », dans Elisabetta COSTA, Patricia LANGE, Nell HAYNES et Jolynna SINANAN (dir.), *The Routledge Companion to Media Anthropology*, Londres, Routledge, p. 505-516, DOI : <https://doi.org/10.4324/9781003175605-49>.

ZIELINSKA Anna (2022), « “Nous avons besoin d’une conversation sur la sexualité”. Politiques de la sexualité : autour de la notion de queer », *Revue française d'éthique appliquée*, 13(2), p. 17-20, DOI : <https://doi.org/10.3917/rfeap.013.0017>.

ZWER Nephthys (2024), *Pour un spatio-féminisme. De l'espace à la carte*, Paris, La Découverte.

NOTES

1. Couldry et Mejas (2019) définissent le colonialisme des données comme l'appropriation de la vie humaine par sa conversion et décomposition en données selon une logique capitaliste. L'expression sera reprise par Ricaurte (2019) pour parler d'un régime des données servant la colonialité du pouvoir, à savoir la domination matérielle et symbolique de l'Europe sur les autres peuples (Quijano, 2007).
2. L'expérience est ici considérée comme « un complexe de significations, d'effets, d'habitudes, de dispositions, d'associations et de perceptions qui résultent de l'interaction sémiotique entre le soi et le monde extérieur » (De Lauretis, 2023, p. 71-72).
3. Traduction de l'auteur. Texte original : « A heuristic, polysemic tool to think politically about big data from the perspective of users and citizens, analysing in particular the contemporary evolution of activism vis-à-vis datafication ».
4. Traduction de l'auteur. Texte original : « The ubiquity of digital communication and mediation dynamics in the platform society as well as on the nature of contemporary activism and its evolving protest/media configurations ».

5. Traduction de l'auteur. Texte original : « ...making strangers into subjects, those who in being included are also willing to consent to the terms of inclusion ».
6. Pour Monro (2019), les personnes non-binaires et *gender queer* (NBGQ) ne s'identifient pas selon un genre exclusivement masculin ou féminin. Leur identité de genre est située au-delà de la binarité de genre, fluctuant entre les genres ou rejetant la binarité de genre.
7. Voir : <https://www.reuters.com/world/us/us-health-agencies-scrub-hiv-other-data-remove-gender-ideology-2025-02-01/>.
8. Voir : <https://www.wired.com/story/artificial-intelligence-lgbtq-representation-openai-sora/>.
9. *Grindr* est une application de rencontres pour des personnes gays, bisexuelles, bicurieuses ou encore *queer*.
10. Voir : <https://www.tf1info.fr/international/egypte-quand-la-police-traque-les-homosexuels-avec-grindr-1557892.html>, <https://www.theguardian.com/global-development/2025/oct/10/torture-blackmail-extortion-the-dangers-of-queer-online-dating-in-ghana>.
11. Voir : <https://www.abc.net.au/news/2025-07-06/gay-dating-app-users-lured-into-violent-homophobic-attacks/105464048>, https://www.lemonde.fr/en/france/article/2024/05/26/homophobic-ambushers-are-baiting-and-beating-gay-men-across-france_6672738_7.html.
12. Traduction de l'auteur. Texte original : « This involves the use of "queer" as a verb to challenge and resist expectations and norms that currently surround data and datafication ».
13. Voir : <https://bigtata.org/>.
14. Voir : <https://fr.homosaurus.org/>.
15. Traduction de l'auteur. Texte original : « Queer is not a category but free movement in the residual space between categories ».
16. Traduction de l'auteur. Texte original : « ...redeployed, twisted, queered from a prior usage and in the direction of urgent and expanding political purposes ».
17. Traduction de l'auteur. Texte original : « Gender cannot be performed independently of the sociotechnical classification structures that humans set on computers and computers expect from humans ».
18. Nous reprenons ici les caractéristiques de la littérature critique des données telle que proposée par Fotopoulou (2020).
19. Voir : <https://www.merriam-webster.com/dictionary/literacy#dictionary-entry-1>.
20. Traduction de l'auteur. Texte original : « only constellations of forms and degrees of literacy ».
21. Le système est défini par Ellul et Porquet (2012) comme un ensemble d'éléments en relation les uns avec les autres de telle façon que toute évolution de l'un provoque une évolution de l'ensemble, toute modification de l'ensemble se répercutant sur chaque élément.
22. Traduction de l'auteur. Texte original : « ...locates, surfaces, and embeds the historical, social, political, and value contexts of the knowing agents — be they human, machinic, or cybernetic — into the representational and operational processes of datafication ».
23. Le *pharmakon* est employé par Stiegler (2008) pour désigner à la fois le poison et le remède. Dans le cas des données, elles apparaissent ainsi soit comme un poison dans les cas de la surveillance des individus, soit comme un remède en permettant d'aiguiller les décisions d'une organisation.
24. Traduction de l'auteur. Texte original : « AI supercharges the risks of datafication by making categorisation and sorting simultaneously more dynamic and more probabilistic, and by integrating these capacities into the backgrounds of everyday systems — that is, part of (digital) infrastructures ».
25. Traduction de l'auteur. Texte original : « Data visceralizations are representations of information rely on multiple senses including touch, smell, and even taste, working together to stimulate our feelings as well as our thoughts ».

26. Traduction de l'auteur. Texte original : « The term "somatechnologies" refers to somatechnics "used" to highlight the inextricability of soma and techne, of the body (as a culturally intelligible construct) and the techniques (dispositifs and "hard technologies") in and through which bodies are formed and transformed ».

27. Ceci fait écho aux recherches d'Haraway *et al.* (2007) où la science n'est pas perçue comme neutre, mais est un domaine de coercition.

RÉSUMÉS

Des pratiques généralistes de traitement des données font subir aux personnes *queer* une invisibilisation de leurs quotidiens, une essentialisation de leurs caractéristiques et des stigmatisations du fait de mauvais calibrage des jeux de données et des algorithmes associés. L'objectif de cet article est de poser les bases d'une littératie des données *queer* afin de combattre les effets négatifs de la *datafication* et proposer des formes alternatives de données pour mieux représenter les vies *queer*.

As a result of certain datafication practices, data essentialization causes queer people to suffer invisibility in their daily lives and discrimination due to poorly calibrated datasets and associated algorithms. The aim of this article is to lay the foundations for queer data literacy in order to challenge the negative effects of datafication and propose alternative forms of data to better represent queer lives.

Las prácticas generalizadas en el procesamiento de datos hacen padecer a las personas *queer* una invisibilización de su cotidiano, esencialización de sus características y estigmatización debido a la clasificación imprecisa del conjunto de datos y algoritmos asociados. El objetivo de este artículo es sentar las bases para los escritos de la literatura de datos *queer* con el fin de contrarrestar los efectos negativos de la dataficación y proponer formas alternativas de los datos para representar mejor las vidas *queer*.

INDEX

Mots-clés : datafication, littératie des données queer, activisme des données, justice des données, violence des données

Palabras claves : activismo de datos, dataficación, justicia de datos, alfabetización de datos, violencia de los datos

Keywords : data activism, data justice, data violence, datafication, queer data literacy

AUTEUR

UGO VERDI

Ugo Verdi est Maître de conférences en Sciences de l'information et de la communication à l'université de Montpellier-Paul-Valéry. Courriel : ugo.verdi@univ-montp3.fr

L'artisanat des données à l'ère de la *datafication*

François Lambotte et Kevin Carillon

- 1 Les discours et les imaginaires associés à la *datafication* de la société, c'est-à-dire la propension à considérer chaque aspect du réel comme une source potentielle de données et, partant, de valeur économique (Mayer-Schönberger et Cukier, 2013), mettent en permanence l'accent sur la dimension quantitative des données. Que l'on parle des *big data* ou des immenses masses de données ayant servi à l'entraînement des IA génératives, tout semble concourir à l'idée que nos régimes contemporains de production et d'exploitation de données se suffisent d'une accumulation toujours plus étendue et intensive pour exister. Pour autant, les spécialistes de la donnée, comme la littérature scientifique, ont depuis longtemps battu en brèche cette idée en mettant en avant la qualité des données comme un enjeu crucial (Cai et Zhu, 2015 ; Dimas, Goldkind et Konrad, 2023). Dans ce contexte, une donnée pertinente est donc une donnée sélectionnée, altérée ou façonnée. Dès lors, loin de l'idée véhiculée par l'appellation « donnée brute » (celle d'une donnée immanente, disponible et dont il suffirait de se saisir), les données qui soutiennent ces régimes apparaissent comme des construits (Gitelman, 2013 ; Barrowman, 2018).
- 2 Or, qui dit construit dit également travail de construction. Dans ce contexte, il est symptomatique de constater qu'une grande partie de ce travail est largement invisibilisée (Denis, 2018 ; Gray et Suri, 2019 ; Rothschild *et al.*, 2022). Invisibilisée, tout d'abord, par les longues chaînes de médiations sociotechniques qui composent les assemblages de données, dont les composants sont dispersés dans le temps et l'espace. Invisibilisée, ensuite, par la nature souvent répétitive, laborieuse et peu valorisée des tâches assurées par des travailleureuses humaines permettant la production des données. Invisibilisée, enfin, par la recherche elle-même, qui s'intéresse davantage au centre des régimes de production et d'exploitation (Muller *et al.*, 2019 ; Stelmaszak, 2022) qu'aux marges et aux périphéries au sein desquelles ces tâches ont lieu.
- 3 Cet article vise à mettre en lumière et à caractériser une partie de ce travail de construction en le saisissant à travers le concept d'artisanat des données (*data crafting*). Relevant du champ plus large du travail des données, entendu ici comme l'ensemble

des activités liées « à la création, la collecte, la gestion, la conservation, l'analyse, l'interprétation et la communication des données » (Bossen *et al.*, 2019, p. 466), l'artisanat des données permet de désigner un ensemble de pratiques qui échappent aux pures logiques d'exploitation ou de précarisation (Casilli, 2015 ; Irani, 2015 ; Gray et Suri, *op. cit.*) dans le travail des données. Ainsi met-il davantage l'accent sur les savoir-faire, le soin apporté aux données et les communautés de pratique qui façonnent les données dans des contextes situés. Autrement dit, là où la littérature a surtout décrit la face sombre du travail des données, le concept d'artisanat ouvre la possibilité de penser des pratiques qui sont certes laborieuses et souvent invisibilisées, mais qui mobilisent également des formes d'expertise, d'engagement collectif et de sens pour les travailleur·euse·s concerné·es.

- 4 Pour développer notre proposition, nous proposons, tout d'abord, de passer en revue la manière dont la littérature scientifique caractérise l'artisanat, puis de dresser un panorama des études sur le travail des données afin de mettre en évidence l'absence d'une perspective proprement artisanale dans ces travaux. En vue d'illustrer ce que nous entendons précisément par artisanat des données, nous mobiliserons ensuite deux terrains ethnographiques. Le premier est une étude menée dans des organismes responsables de la production de registres du cancer (RC) dans deux départements français, tandis que le second concerne la communauté Stack Overflow (SO) sur Stack Exchange, un forum en ligne fondé par des développeurs de logiciels, devenu une référence dans le domaine de la programmation. À travers ces données empiriques, nous montrerons en quoi le travail des acteur·rice·s de ces deux terrains s'apparente à de l'artisanat des données tout en en mettant en exergue ses caractéristiques. Enfin, nous analyserons la manière dont les artisan·es des données s'appuient, précisément, sur le caractère artisanal de leurs pratiques pour faire sens de la massification des données et des potentialités de l'IA dans leurs domaines respectifs.
- 5 Cet article contribue ainsi à la littérature sur le travail des données à l'ère de la *datafication*. D'abord, il propose un concept théorique original pour penser et analyser un ensemble de pratiques qui s'inscrivent dans certains contextes de production. En mettant en valeur le travail des artisan·es de la donnée, il invite également à une prise en compte accrue des tâches invisibles, et pourtant essentielles, qui participent au fonctionnement et à la persistance de nos régimes de production et d'exploitation des données. Enfin, en soulignant les compétences, les savoir-faire et les valeurs qui sous-tendent ces pratiques, l'artisanat des données met en évidence la dimension à la fois éthique et politique de certaines formes de travail des données, souvent occultée par les discours technocentrés sur la *datafication* et les progrès de l'IA.

L'artisanat au croisement du développement d'habiletés, du *care* et de la matière

- 6 Comment la littérature scientifique appréhende-t-elle l'artisanat ? Cette question constitue le point de départ de la présente section, qui explore les principales dimensions mises en évidence par la littérature pour caractériser ce phénomène. Les recherches sur l'artisanat s'articulent ainsi autour de trois dimensions complémentaires : (1) le développement d'habiletés au sein de communautés de pratique, (2) le soin apporté aux choses et (3) le rapport à la matière travaillée. En croisant ces trois dimensions, la littérature offre une compréhension fine de l'artisanat

par rapport à d'autres types de pratiques, tout en ouvrant la voie à des réflexions sur ses évolutions face aux défis contemporains tels que (4) l'automatisation et la numérisation.

Habiletés et communautés de pratique

- 7 L'artisanat, tel que le conçoit Sennett (2008), est à la fois une pratique et une éthique dont l'objectif principal est de « faire un bon travail pour lui-même ». Ici, la qualité prévaut sur l'efficacité, s'imposant comme la valeur cardinale (Kolko, 2011). Il s'agit d'un processus visant à créer un artefact fonctionnel, dans lequel conception et fabrication sont étroitement imbriquées (Kolko, citant Wroblewski, *ibid.*). Cette imbrication du design et de la fabrication en fait une pratique de cognition située, impliquant une « grande réflexivité sur les habiletés, les outils, les méthodes et les matériaux à mobiliser pour atteindre un objectif de faire les choses bien » (Schwalbe, 2010, p. 109, notre traduction).
- 8 L'artisanat repose ainsi sur le développement d'habiletés, fruit d'un apprentissage par essais et erreurs. Il exige un entraînement et un travail centré sur une tâche précise, parfois très simple, mais répétée continuellement (Kolko, *op. cit.*). Cette répétition permet à l'artisan·e d'acquérir une connaissance approfondie de son métier et de développer sa réflexivité. Comme l'illustre le travail de Portisch, cité par Ingold (2019), sur les artisanes en Mongolie, cette réflexivité et cette évaluation continues sont au cœur de la pratique, tant pour les novices que pour les artisanes expérimentées, et ce, à chaque étape du processus. L'acquisition et le développement des habiletés dans l'artisanat se fondent sur le principe de maître-apprenti·e et sur l'existence d'une communauté de pratique, définie comme le regroupement d'individus engagés dans une activité commune, quel que soit leur niveau d'expertise, et qui se caractérise par une structure évolutive où coexistent des membres expérimentés et des novices (Lave et Wenger, 1991). Aussi la pratique de l'artisanat et les valeurs qui la sous-tendent impliquent-elles que des personnes plus expérimentées partagent leur savoir avec d'autres personnes, nouvellement venues. Selon Klammer (2012, p. 7), puisque la valeur centrale de l'artisanat est la qualité, on désignera comme « maître » des personnes ayant la capacité de juger et de défendre la qualité dans la pratique, dont les habiletés et la créativité sont reconnues par les pairs et qui possèdent des qualités pédagogiques permettant la transmission du savoir. L'apprentissage, cependant, ne se limite pas à une simple observation passive. Il repose sur l'implication active de l'apprenti·e qui, au fil de son travail, interagit avec les artisan·es aguerri·es, observe leurs gestes et acquiert progressivement ses propres habiletés (Fyhn et Søraa, 2017). Cette relation favorise un apprentissage par la pratique, où la répétition des gestes s'accompagne d'une réflexion critique sur la qualité du travail accompli. Cette communauté de pratique permet alors non seulement l'acquisition progressive des compétences, mais renforce aussi l'ancrage des valeurs fondamentales de l'artisanat telles que la quête de qualité et le souci du travail bien fait (Kolko, *op. cit.*). La communauté de pratique joue ainsi un rôle essentiel dans la préservation, la transmission et la perpétuation de ces valeurs.

Care des choses et politique du faire

- 9 Cette quête de la qualité et ce souci du travail bien fait permettent de rattacher la pratique artisanale à une forme de *care*, entendu ici dans une perspective postmoderne. Le *care*, que l'on peut traduire par le fait de « s'occuper de », « faire attention », « prendre soin » ou encore « se soucier de » (Laugier et Molinier, 2009), a d'abord émergé dans la philosophie morale et les études féministes comme une manière de repenser l'éthique des relations interpersonnelles. Il s'agissait alors de mettre en valeur une éthique située, attentive aux vulnérabilités et aux interdépendances constitutives des relations humaines, tout en révélant la dimension genrée (c'est-à-dire presque exclusivement féminine) et souvent invisibilisée du travail du soin. Le concept a néanmoins progressivement migré vers d'autres disciplines, notamment les *science and technology studies* (STS), où il a été reformulé dans une perspective élargie aux non-humains.
- 10 Détaché de son ancrage exclusivement interpersonnel et genré, le *care* s'est vu appliqué aux objets, aux environnements et aux infrastructures, ouvrant la possibilité de penser des relations de soin qui ne concernent pas uniquement les humains, mais également les non-humains. Bruno Latour (2004), avec sa distinction entre *matters of fact* et *matters of concern*, a ouvert la voie en invitant à considérer les faits techniques et scientifiques comme des réalités toujours déjà enchâssées dans des préoccupations, des controverses et des attachements. Maria Puig de la Bellacasa (2011) a prolongé cette réflexion en introduisant la notion de *matters of care*, insistant sur le fait que ces préoccupations ne sont pas uniquement cognitives ou théoriques, mais impliquent également une dimension affective et éthique envers les non-humains.
- 11 Ce « soin des choses » comporte également une dimension politique (Laugier et Molinier, *op. cit.* ; Martin, Myers et Viseu, 2015). Prendre soin des choses, c'est non seulement garantir leur bon fonctionnement, mais aussi interroger les logiques qui président à leur production, à leur circulation, à leur persistance et à leur valorisation. L'artisanat se situe précisément dans cet espace. En réintroduisant du temps, de l'attention et de la minutie dans la production, il incarne une résistance discrète, mais structurante à l'hégémonie des logiques industrielles. C'est pourquoi plusieurs travaux soulignent la proximité structurelle entre artisanat et *care* (Kunneman, 2013 ; MacGill, 2019). Vasiliki Tsaknaki et Anna Vallgård (2023) rappellent ainsi que l'artisanat est fondamentalement « a matter of care ». Dans cette perspective, le soin et l'attention ne constituent pas des composantes secondaires, mais des principes constitutifs de la pratique artisanale. L'artisan·e ne se contente pas de produire un objet fonctionnel ; il ou elle exprime une forme d'engagement envers la matière travaillée.

L'artisanat comme rapport particulier à la matière

- 12 Certain·es auteur·rice·s, comme Klammer (*op. cit.*), défendent une conception rigoureuse de l'artisanat, centrée exclusivement sur le travail manuel et nécessitant des compétences acquises par l'entraînement dans le cadre d'une pratique soutenue. Dans cette perspective, un travail mobilisant des savoir-faire techniques et des habiletés spécifiques, mais où le geste manuel n'est pas central, ne saurait être considéré comme relevant, à proprement parler, de l'artisanat. Pour d'autres tels que Kuijpers (2019), l'essence de l'artisanat réside dans l'interaction dynamique entre l'artisan·e et la

matière. Selon ce dernier, « l'habileté, en tant que processus récursif et perceptif, crée des connaissances à partir des interactions matérielles » (p. 264, notre traduction). Ce n'est donc pas seulement la finalité de l'objet produit ou le geste qui importent, mais le processus même par lequel l'artisan·e développe sa compréhension du matériau à travers l'action. La matière façonne tout autant la pensée que la main : la pratique artisanale devient alors un dialogue entre l'artisan·e et le matériau.

- 13 L'exemple de l'âge du bronze illustre bien cette idée. Kuijpers (*ibid.*) montre que les propriétés spécifiques du bronze (sa malléabilité, sa durabilité et sa capacité à être fondu et remodelé) ont influencé les pratiques artisanales, facilitant l'émergence de techniques comme le design, le prototypage et la reproduction à grande échelle. Ces innovations ont profondément transformé les dynamiques sociales et économiques de l'époque, au point de donner naissance à une nouvelle ère.
- 14 L'artisan·e, dans ce processus, développe une connaissance intime du matériau par la compréhension de ses propriétés, de ses contraintes et de son histoire. Cette connaissance va au-delà d'une simple maîtrise technique, dans la mesure où elle intègre une dimension créative et réflexive qui permet non seulement de façonner la matière, mais aussi de l'altérer et de l'adapter à de nouveaux usages (Shorter, 2015, p. 4). L'artisan·e devient ainsi un·e médiateur·rice entre la matière brute et la forme achevée, inscrivant son geste dans une continuité historique et culturelle.

L'artisanat à l'ère des machines et de la numérisation

- 15 L'industrialisation et l'arrivée des machines dans le processus de fabrication ont provoqué plusieurs débats scientifiques quant à la notion d'artisanat en lien avec ce rapport à la matière (Fyhn et Søraa, *op. cit.* ; Kolko, *op. cit.* ; Schwalbe, *op. cit.*). Pour Ingold (*op. cit.*), la distinction fondamentale entre artisan·e et machiniste réside précisément dans leur relation à la matière. Cette réflexion s'appuie sur les travaux de Pye (1995), qui interroge la notion d'artisanat en la confrontant aux logiques de fabrication industrielle, notamment la production de masse. Selon lui, il est réducteur, voire inutile, d'opposer strictement le travail manuel au travail assisté par des machines ou des outils. Il propose plutôt de distinguer deux modes de fabrication : la fabrication à risque et la fabrication certaine.
- 16 La fabrication à risque implique une interaction continue entre l'artisan·e et la matière. Le résultat final dépend étroitement des compétences, de la dextérité et du jugement de l'artisan·e. Fyhn et Søraa illustrent cette idée en prenant l'exemple du menuisier : « Durant tout le processus, le menuisier doit rester alerte et pleinement engagé, en interaction directe avec la matière, car le résultat est en permanence à risque » (*op. cit.*, p. 75, notre traduction). L'instabilité propre à la création à risque crée ainsi un espace propice à l'expérimentation, à l'innovation et à la créativité. En revanche, dans le cas de la fabrication certaine, typique des environnements industriels, les machines sont configurées pour produire des résultats prévisibles et standardisés, sans considération pour le jugement ou la dextérité de l'opérateur·rice. Par exemple, si le même menuisier utilise une machine-outil pour découper les pièces d'une armoire en usine, il n'interagit plus directement avec la matière : il devient opérateur de la machine, perdant ainsi le lien sensible avec le matériau brut. Pye ne dénie pas pour autant les avantages de la fabrication certaine, particulièrement dans les contextes de production à grande échelle. Toutefois, il invite à réfléchir sur les pertes engendrées par l'effacement du

risque dans le processus. Schwalbe souligne également l'existence de ces pertes potentielles dans le cadre de la fabrication certaine, notamment « l'habitude de prendre soin dans le processus de travail, la dextérité et le jugement nécessaires à la confection manuelle, l'appréciation de la finesse du travail réalisé sans l'aide de la machine, les occasions d'innovation imprévue, la réflexion sur les limites de nos habiletés ainsi que l'esthétique de la diversité » (*op. cit.*, p. 110, notre traduction).

- 17 Traditionnellement, les recherches sur l'artisanat se sont concentrées sur la manipulation de matériaux physiques et sur la relation sensible entre l'artisan-e et la matière. Cependant, ces dernières années, un nouveau champ d'études s'est développé autour de l'artisanat numérique, qui interroge le rôle des outils numériques dans le processus créatif comme la production d'artefacts numériques (Treadaway, 2007 ; Zoran et Buechley, 2013 ; Grimshaw, 2017 ; Bernabei et Power, 2018). Cette approche s'intéresse à la manière dont les technologies numériques peuvent s'inscrire dans une logique artisanale en préservant des principes tels que la maîtrise technique et la sensibilité au détail lorsque la matière travaillée n'est plus physique.
- 18 Plusieurs chercheur-euse-s ont notamment élargi la notion d'artisanat aux pratiques numériques. Ces dernier-e-s soutiennent, par exemple, que la manipulation du clavier pour écrire du code peut être assimilée à un travail manuel, dans la mesure où elle mobilise des compétences techniques, de la créativité et un rapport direct à la « matière numérique » (Nachtigall, 2019). Dans le domaine du développement logiciel, cela s'est traduit par l'émergence du mouvement autour du *Manifeste pour l'artisanat du logiciel*. Né en réaction aux limites perçues des méthodes agiles, ce mouvement promeut une vision élargie de l'artisanat, centrée sur la qualité du code, la transparence, le sens des responsabilités ainsi que l'importance des dynamiques communautaires et du jugement des pairs (Sundelin *et al.*, 2021).
- 19 Finalement, une question persiste : les données sont-elles simplement au service de l'artisanat ou peuvent-elles devenir le matériau même de la création ? Cette question devient cruciale lorsqu'on se penche sur des formes d'artisanat où la matière (les données) se réduit à des éléments considérés comme abstraits — à savoir des 1 et des 0 —, comme c'est le cas dans la programmation ou la gestion de données (Shorter, *op. cit.*). Dans sa thèse consacrée aux liens entre artisanat et données dans le design industriel, Nachtigall (*op. cit.*) explore la notion d'artisanat des données et s'interroge sur leur matérialité. En s'appuyant sur les travaux de Dourish (2017), il souligne que les données, bien que souvent perçues comme immatérielles, reposent sur des infrastructures tangibles (code, systèmes, protocoles, machines) qui conditionnent la manière dont elles sont collectées, manipulées et visualisées. Cette matérialité confère aux données une forme de consistance qui les rend « façonnables » (Lambotte et Carillon, 2024). Nachtigall va au-delà de la question de la matérialité des données et s'intéresse davantage aux caractéristiques qui définissent l'artisanat des données en tant que pratique : l'apprentissage par la pratique, l'appartenance à une communauté, la maîtrise d'outils spécifiques et la place du jugement humain dans le processus de création. Selon lui, c'est par ces dimensions que l'on peut légitimement parler d'un artisanat des données, même lorsque la matière travaillée est numérique.
- 20 Notre approche s'inscrit ainsi dans la lignée des travaux de Nachtigall et propose de considérer l'artisanat des données comme un cadre d'analyse critique des processus de construction et d'exploitation des données. Dans ce contexte, nous considérons les données comme sociomatérielles, au sens où leur existence repose sur un ensemble

d'infrastructures qui participent de leur matérialité tangible (Bates, Lin et Goodale, 2016), mais aussi au sens de *matter* (Cooren, 2015), à savoir de leur poids dans les réalités sociales qu'elles constituent. Enfin, nous considérons l'artisanat des données comme une production « à risque », à savoir comme un espace propice à la réflexivité, à la créativité et à l'innovation et dont la persistance est remise en question par l'automatisation et la standardisation.

L'absence d'une perspective artisanale sur le travail des données

- 21 Avec l'intensification des processus de *datafication* (Mayer-Schönberger et Cukier, *op. cit.*), la question du travail des données (*data work*) s'est imposée comme un enjeu central dans la littérature scientifique. Dans la présente section, nous inscrivons notre proposition d'artisanat des données dans les études consacrées au travail des données. Cette exploration de la littérature sera l'occasion de montrer que, bien que certaines caractéristiques du travail artisanal soient parfois évoquées, aucun de ces travaux ne propose une conception théorique globale et cohérente permettant d'appréhender l'ensemble des dimensions spécifiques au travail artisanal des données.

Situer l'artisanat des données dans les études sur le travail des données

- 22 Le travail des données est défini par Bossen *et al.* comme toute activité « liée à la création, la collecte, la gestion, la conservation, l'analyse, l'interprétation et la communication des données » (*op. cit.*, p. 466, notre traduction). Englobant une vaste gamme d'activités et de pratiques (allant de l'encodage au nettoyage des données, en passant par leur standardisation, leur contextualisation ou la maintenance des bases), ce travail a longtemps été appréhendé par la littérature scientifique comme un « processus rationnel, guidé par les données, de “découverte” » (Rothschild *et al.*, *op. cit.*, p. 5, citant Muller *et al.*, *op. cit.*, notre traduction), qui relevait du ressort d'experts ou de spécialistes (Carter et Sholler, 2015). Cependant, comme le soulignent Rothschild *et al.* (*op. cit.*), la littérature en sciences des données et plus largement celle consacrée au travail des données ont amorcé, ces dernières années, deux décentrement successifs visant à dépasser ce cadrage restrictif. Notre proposition d'artisanat des données s'inscrit dans ces déplacements, tout en dialoguant avec les diagnostics critiques sur l'exploitation des formes invisibles de travail des données.
- 23 Le premier décentrement a consisté à déplacer l'attention des procédures techniques vers les pratiques situées et incarnées qui sous-tendent le travail des données. Ce déplacement a progressivement permis de déconstruire l'illusion d'objectivité et de rationalité attachée aux données, pour montrer que leur production est façonnée par des dynamiques sociales, cognitives et organisationnelles. Dans ce contexte, Koesten *et al.* (2021) identifient des schémas d'activités récurrents dans l'exploration et l'interprétation des données, révélant que ces tâches mobilisent des stratégies cognitives et des interactions sociales autant que des compétences techniques. De leur côté, D'Ignazio et Klein (2023) insistent sur la réflexivité nécessaire pour contextualiser et redonner sens aux données, tandis que Miceli, Schuessler et Yang (2020) mettent en évidence l'inscription de ces processus dans des rapports de pouvoir, où valeurs,

intérêts et priorités façonnent les résultats, en générant des frictions (Bates, Lin et Goodale, *op. cit.*).

- 24 Notre proposition d'artisanat des données s'inscrit pleinement dans ce décentrement. Le recours à l'artisanat permet, en effet, d'insister sur le caractère situé et incarné (Choroszewicz, 2022) de certaines pratiques en matière de production de données, dans la mesure où ces pratiques reposent sur des formes de jugement et d'expertise qui échappent aux représentations strictement techniques du travail des données. En effet, penser le travail des données en termes artisanaux revient à souligner que les données sont travaillées dans des contextes particuliers, mobilisent des expertises qui ne sont pas toujours codifiées et nécessitent des arbitrages continus qui associent savoirs expérimentiels et inscription dans des collectifs.
- 25 Le deuxième décentrement s'est opéré par l'élargissement de la focale au-delà des expert·es pour inclure une pluralité d'acteur·rice·s intervenant dans la production et l'usage des données. Foster *et al.* (2018) proposent, par exemple, d'analyser le travail des données à travers trois niveaux (macro, méso et micro) et montrent que de nombreux profils, aux statuts et compétences variés, participent à la construction, à l'interprétation et à la circulation des données. En mettant en avant la figure de l'artisan·e et son inscription dans des communautés de pratique, l'approche artisanale prolonge ce déplacement en fournissant une manière de décrire concrètement comment cette pluralité s'organise dans certains contextes de production. Elle met en évidence la dimension collective du travail des données, fondée sur des communautés, des routines partagées et des dynamiques d'apprentissage qui assurent la cohésion et la continuité de ces pratiques.
- 26 Ce deuxième décentrement a également conduit la recherche à porter attention aux formes de travail périphériques (Rothschild *et al.*, *op. cit.*) qui soutiennent les infrastructures de données, c'est-à-dire aux activités des travailleuse·s de l'ombre « qui n'ont pas le rôle d'acteur, mais qui, par leur travail technique, permettent que la pièce se joue et que le public la voie » (Waquet, 2022, p. 8). C'est dans cette perspective que s'inscrit la littérature sur le *digital labor*, qui étudie les nouvelles formes de travail associées aux plateformes numériques et aux infrastructures de données, en mettant en évidence des dynamiques d'exploitation, de précarisation et de fragmentation (Casilli, *op. cit.*). Cette littérature couvre un spectre allant du micro-travail rémunéré sur des plateformes comme Amazon Mechanical Turk (où des tâches répétitives et peu qualifiées sont effectuées par des travailleuse·s dispersée·s et sous-payée·s) aux contributions non rémunérées des utilisateur·rice·s, dont les données sont captées et monétisées par les grandes entreprises du numérique. Gray et Suri (*op. cit.*) parlent ainsi de « travail fantôme » (*ghost work*) pour désigner ces tâches invisibles d'annotation, de nettoyage et d'organisation des données, tandis qu'Irani (*op. cit.*), à travers le « micro-travail » (*microwork*), insiste sur l'importance des modérateur·rice·s, tagueur·euse·s et classificateur·rice·s, indispensables au fonctionnement des plateformes, mais relégué·e·s à des positions marginales.
- 27 L'artisanat des données n'a pas pour vocation de minimiser la critique des rapports d'exploitation mise en évidence par le *digital labor*. Il s'agit plutôt d'un éclairage supplémentaire sur des pratiques concrètes qui, dans certains contextes de production et en fonction des finalités poursuivies, participent à conférer sens et valeur aux données. Cette approche permet de souligner que toutes les activités de production de données ne se réduisent pas aux seuls régimes d'exploitation et de précarisation,

notamment dans les contextes où le travail des données ne repose pas prioritairement sur l'extraction d'une plus-value économique immédiate, mais sur des finalités collectives, sociales ou épistémiques. Pour autant, ces formes de contribution demeurent insuffisamment reconnues ou valorisées et, partant, leur rôle dans le maintien des infrastructures de données reste largement invisibilisé (Denis, *op. cit.*).

Des « traits artisanaux » dispersés dans la littérature

- 28 Des « traits artisanaux » apparaissent de manière récurrente dans les travaux sur le travail des données, mais toujours de façon partielle, dispersée et sans donner lieu à une théorisation cohérente. Par traits artisanaux, nous entendons l'une des trois dimensions constitutives de l'artisanat identifiées plus haut, qu'il s'agisse d'une pratique orientée vers la qualité des données et inscrite dans des logiques de *care*, du développement d'habiletés collectives autour des données au sein de communautés de pratique ou encore d'un rapport « à risque » dans la production des données.
- 29 Les travaux sur les infrastructures de données, notamment, permettent d'entrevoir plusieurs de ces traits. Dans la lignée de Star (1999), l'objectif des *infrastructures studies* était, dès l'origine, d'attirer l'attention sur les « tâches concrètes » et les « ressources matérielles et symboliques » (Beltrame et Peerbaye, 2018, p. 179) qui assurent le maintien des infrastructures. Ces travaux mettent en évidence des pratiques qui vont de la réparation aux opérations de maintenance, en passant par des formes de bricolage situé (Harper, 1987) ou d'improvisation (Henke, 1999 ; Schubert, 2018). Appliquée aux infrastructures de données, cette perspective conduit Denis (*op. cit.*) à décrire ces activités comme un soin particulier porté à nos « infrastructures scripturales ». Dans ce contexte, Dagiral et Peerbaye (2022) mettent ainsi en évidence l'invisibilisation du travail effectué en amont de la collecte de données sur les maladies rares, soulignant la dépendance des infrastructures de données à des efforts méthodiques et minutieux réalisés par des acteur·rice·s souvent relégué·e·s à l'arrière-plan. De manière complémentaire, Stiefel (2018) prolonge cette réflexion en s'intéressant aux « agents opérationnels » dans la gestion des biobanques hospitalières. À travers une approche centrée sur le *care*, elle montre comment ces travailleuse·s assurent la continuité et la fiabilité des infrastructures de données biomédicales ainsi que les affects qui découlent de ce travail.
- 30 Ce lien entre travail des données et *care* a été exploré plus explicitement dans diverses publications (Zakharova et Jarke, 2024). Tandis que Jarke et Büchner proposent le concept de *data care arrangements* pour répondre à la question « pourquoi et comment les personnes dans les organisations se soucient des données » (2024, p. 702), Pinel, Prainsack et McKevitt montrent, dans leur étude de cas, que « les chercheurs établissent des relations à long terme avec les données qu'ils contribuent à produire et se sentent responsables de leur épanouissement et de leur développement ». Ils soulignent également que « ces pratiques de soin [...] contribuent à rendre les données précieuses pour l'institution » (2020, p. 175). Davies et Holmer (2024) étudient des dynamiques similaires dans le monde universitaire en mettant l'accent sur le fait que ces pratiques de *care* envers les données peuvent être lues comme des formes de résistance au discours dominant sur l'excellence universitaire. Enfin, Baker et Karasti (2018) approfondissent la dimension éthique et politique du *care* en mettant en lumière la manière dont le travail des données à l'échelle locale, dans le cadre du déploiement

de grandes infrastructures, est souvent relégué au rang de « chose négligée » et elles analysent les pratiques qui émergent en réponse à cette tension.

- 31 D'autres travaux mettent davantage l'accent sur les collectifs et les apprentissages dans le travail des données. Rothschild *et al.* (*op. cit.*), en mobilisant de manière explicite le concept de communauté de pratique, cherchent à dépasser une vision strictement occupationnelle des compétences pour mettre en évidence les dynamiques collectives, les valeurs partagées et les processus d'apprentissage qui structurent le travail des données. De telles dynamiques collectives se retrouvent dans plusieurs autres études (Cragin et Shankar, 2006 ; Benfeldt, Persson et Madsen, 2020). Néanmoins, si ces études traitent des dynamiques collectives dans le travail des données, elles restent cantonnées à l'analyse de la transmission des savoirs et de la socialisation professionnelle, sans les inscrire dans une conceptualisation plus large.
- 32 Pris ensemble, ces travaux montrent que les traits artisanaux sont donc présents dans la littérature, mais de manière dispersée, apparaissant en filigrane plutôt que comme les éléments d'un cadre unifié. C'est précisément cette dispersion que notre proposition vise à dépasser en proposant le concept d'artisanat des données.

L'artisanat des données au cœur de deux terrains ethnographiques

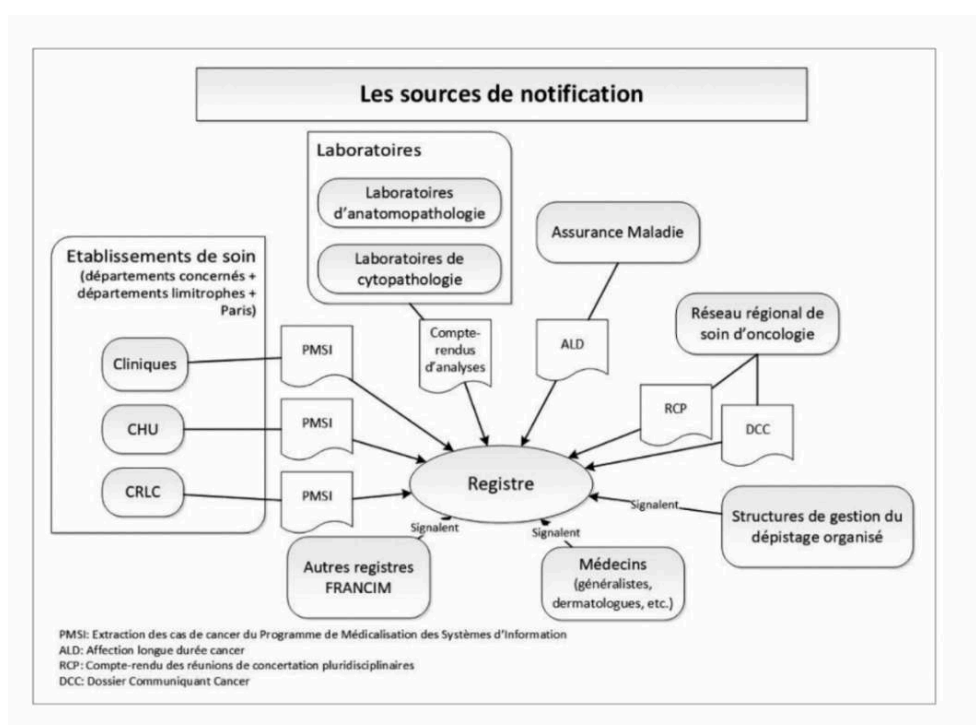
- 33 Si notre contribution est avant tout conceptuelle, elle ne se réduit pas à une abstraction détachée du terrain. Au contraire, nous entendons démontrer que l'artisanat des données s'ancre dans des réalités empiriques et permet, dans le même temps, d'éclairer des pratiques concrètes liées au travail de la donnée. Pour ce faire, nous nous appuyons sur des matériaux issus de deux terrains ethnographiques distincts, menés par l'un des auteurs. Ces données nous permettront d'examiner comment le travail des acteur·rice·s concerné·e·s s'apparente à un véritable artisanat des données, tout en mettant en exergue ses caractéristiques.
- 34 Nous nous appuyons d'abord sur un travail ethnographique (Martin-Scholz, Mayère et Lambotte, 2021) mené dans des organismes responsables de la production de RC dans deux départements français, qui ont pour mission de réaliser « une collecte continue et exhaustive de données nominatives concernant un ou plusieurs événements de santé [en l'espèce, le cancer] dans une population géographiquement définie, à des fins de recherche et de santé publique, par une équipe disposant des compétences appropriées¹ ». Le travail des RC peut être défini comme la transformation et l'assemblage de données initialement produites à des fins thérapeutiques ou économiques en données pouvant servir à des fins scientifiques et de santé publique. Ce travail se déroule dans le cadre d'un parcours des données, défini comme « le mouvement des données depuis leurs sites de production initiaux [vers le RC] dans lequel elles sont traitées, mobilisées et réutilisées » (Leonelli, 2020, p. 9).
- 35 Notre second terrain ethnographique (Renard *et al.*, 2024) est la communauté SO, un vaste réseau de forums en ligne où des communautés se rassemblent autour de sujets spécifiques. Fondée en 2008, la communauté SO est devenue une référence dans le domaine de la programmation. Sur cette plateforme, les utilisateur·rice·s posent des questions, apportent des réponses, votent sur la pertinence des contenus et peuvent les éditer.

- 36 Le point commun entre ces deux terrains de recherche est qu'ils regroupent des travailleur·euse·s de la donnée dont l'objectif est de produire des données de qualité. Dans un cas, ce travail est destiné à la recherche en sciences de la santé et à la lutte contre les cancers et, dans l'autre cas, il participe au domaine des sciences et des sciences appliquées. Ces deux collectifs organisés ont également en commun d'être bousculés par la « matérialisation des imaginaires de l'IA » (Bareis et Katzenbach, 2022), conduisant à l'émergence de controverses sur le rôle et les conséquences de l'IA sur le devenir des données produites et de l'infrastructure qui soutient celles-ci. Pour Pachidi *et al.* (2021), ces moments de controverse sont intéressants d'un point de vue analytique, car ils tendent à rendre visibles ces organisations ou communautés, qui se fondent habituellement dans les infrastructures dont elles assurent la maintenance.
- 37 C'est en travaillant à la compréhension de la transformation du travail de la donnée pour les RC, d'une part, et sur le mouvement de grève lancé par les modérateur·rice·s de la plateforme SO, d'autre part, que nous nous sommes rendu compte de la pratique située qui sous-tend la production des données. Cette prise de conscience repose sur plusieurs éléments : l'attention minutieuse que les acteur·rice·s portent à la donnée elle-même, les exigences de qualité qui lui sont associées, le savoir-faire nécessaire à sa production ainsi que les dynamiques collectives et les schèmes de valeurs qui structurent ces pratiques. C'est en observant ces processus et en interrogeant les activités des travailleur·euse·s de la donnée que l'image de l'artisan·e s'est progressivement imposée à nous. L'objectif des prochaines sections est donc d'aller au-delà d'une simple intuition et d'explorer ce que la conceptualisation de l'artisanat des données peut apporter à la compréhension du travail de la donnée.

Dans la fabrique des registres du cancer

- 38 En France, les RC sont répartis sur l'ensemble du territoire. Ces organisations de petite taille regroupent des enquêteur·rice·s dont l'objectif principal est d'identifier tous les cas de cancer pour une pathologie spécifique (par exemple, le cancer du sein) dans une région donnée. Une fois les cas répertoriés, l'équipe du RC doit rassembler les informations associées, qui sont souvent dispersées dans diverses bases de données (assurance maladie, hôpitaux, laboratoires, etc.) (voir figure 1).

Figure . Cartographie des sources de notification de la fabrique des données du registre du cancer



Source : Martin-Scholz, Mayère et Lambotte (op. cit.).

- 39 Ce processus implique plusieurs étapes cruciales. Les données doivent d'abord être collectées et centralisées, puis vérifiées et harmonisées au regard de la nomenclature internationale en vigueur, afin de garantir leur comparabilité et leur intégration dans le registre. La qualité des données est un principe fondamental de ce travail, reposant à la fois sur leur exactitude et leur exhaustivité. Ce processus de recontextualisation des données (Leonelli, 2016) requiert une expertise multidimensionnelle du fonctionnement des hôpitaux, de l'organisation du dossier patient, mais également des différences dans les usages des nomenclatures utilisées en milieu hospitalier et dans le cadre des registres.
- 40 Les RC, en tant qu'organisations, ne disposent pas d'un accès direct et systématique à toutes les bases de données hospitalières par l'intermédiaire des réseaux sécurisés (VPN). Par conséquent, les enquêteur-rices doivent se rendre sur place, consulter les dossiers physiques pour scanner et identifier les éléments pertinents. Bien que cette tâche puisse paraître simple, elle nécessite en réalité une connaissance approfondie de la structure des dossiers médicaux, des types d'examen à analyser et des éléments cliniques pertinents à extraire.
- 41 Lors de l'encodage, un autre défi émerge : les typologies utilisées pour caractériser le stade d'avancement d'un cancer peuvent varier entre les pratiques des oncologues, d'une part, et les standards internationaux adoptés par les RC, d'autre part. Cette discordance illustre la complexité du travail de recontextualisation, qui mobilise à la fois des connaissances médicales et une compréhension fine des normes administratives et scientifiques encadrant la production des données de santé.

[Codeuse RC] : Avec cela, j'ai mon [guide] de codage pour chaque topographie et quand je fais les poumons, je sors mon dossier sur les poumons et mes [guides] de codage [voir figure 2] parce que le codage est totalement différent d'un organe à

puisqu'elle, elle va collecter toute l'information qu'elle trouve dans l'endroit où elle est [...] En plus, nous, on essaie d'optimiser, [...] on va sur la source principale d'abord et si là, il n'y a pas assez, on va relancer les autres sources [...] Donc s'il y en a une qui revient, elle sait aussi qu'il n'y a pas assez, elle s'en rend compte. Donc elle dit « non, celle-là, tu vois, j'ai trouvé que ça donc il faut demander à machin d'aller dans tel hôpital parce que j'ai pas assez ».

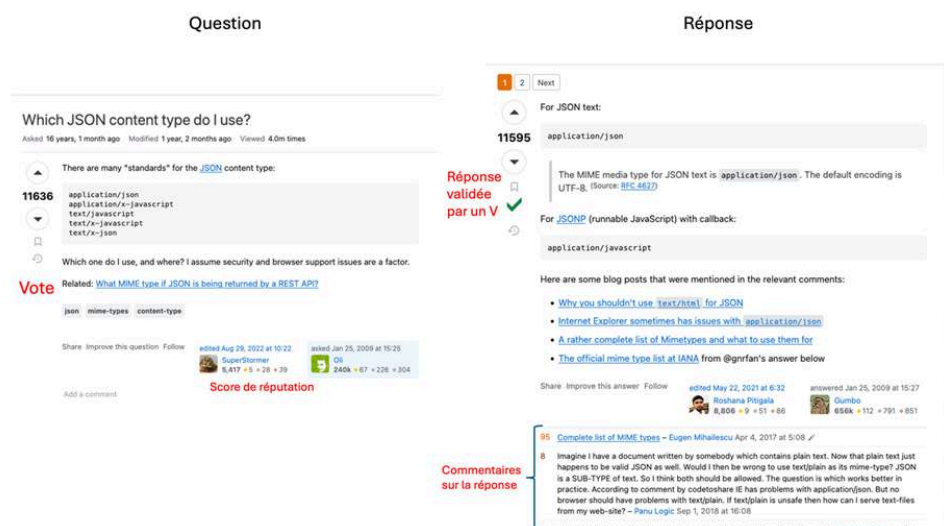
- 44 On constate ainsi que ce travail exige une réflexivité permanente, en particulier face aux avancées de la recherche sur le cancer. Par exemple, le cas d'étude révèle que l'épidémiologie impose une comparabilité historique des données afin d'analyser l'évolution des tendances sur un territoire donné. Cependant, l'amélioration des classifications et la découverte de nouvelles sous-catégories de cancers posent un dilemme : faut-il adapter les registres pour inclure ces distinctions plus fines, au risque de compromettre la continuité des données ? Ces décisions sont débattues et prises à l'échelle collective, au sein de la communauté des RC, représentée par l'association FRANCIM. Cette dernière joue un rôle-clé dans la standardisation des pratiques, en développant des outils logiciels et des guides d'encodage permettant d'harmoniser le travail des registres sur l'ensemble du territoire.
- 45 Le travail des enquêteur·rice·s met également en lumière une approche minutieuse et engagée vis-à-vis de la donnée. Ce soin se manifeste dans les efforts, parfois considérables, déployés pour récupérer des informations essentielles. Par exemple, lorsque l'accès aux données est bloqué par des laboratoires privés, les enquêteur·rice·s vont mettre en place des dispositifs bricolés pour récupérer le stade d'avancement d'une tumeur mentionné en annexe d'un dossier médical. Ce souci du détail dépasse la simple technicité de l'encodage : cet engagement contribue à garantir que les patient·e·s ayant moins de visibilité dans les grandes institutions hospitalières, par exemple, en raison de leur situation socioéconomique, soient correctement recensé·e·s dans les RC.
- [Enquêtrice RC] : Il est évident que si vous passez au big data, vous avez le nombre de cancers ! Mais ensuite, par stade, tout ça... on n'a pas la répartition... après, c'est un gros travail. Quand on parle aux médecins qui nous accueillent... quand ils voient le travail minutieux que nous faisons... parce que souvent, il faut chercher à plusieurs endroits. Je veux dire qu'une machine ne peut pas faire ça !
- 46 À travers l'image de la « machine » présent dans cet extrait, on constate également que les enquêteur·rice·s se positionnent par rapport aux imaginaires sociotechniques liés à l'IA et au *big data* (Bareis et Katzenbach, *op. cit.*). Pour ces dernier·e·s, ces imaginaires véhiculent une vision de la donnée comme une ressource brute, standardisée et immédiatement exploitable, occultant ainsi le contexte de sa production et la complexité de son traitement. Cette vision est problématique dans le cadre des RC, car elle tend à effacer l'importance et le rôle crucial des enquêteur·rice·s ; comme si le travail de sélection, de vérification et de contextualisation de la donnée n'avait pas de valeur. Or, l'automatisation ne peut remplacer l'expertise des enquêteur·rice·s, qui repose sur une compréhension fine des spécificités cliniques, administratives et épidémiologiques. Une approche purement automatisée risquerait de produire, au regard des enquêteur·rice·s, des données incomplètes ou biaisées, compromettant ainsi la fiabilité des analyses en santé publique et in fine des politiques de santé publique qu'elles alimentent. Pour ces artisan·e·s, il est donc essentiel de reconnaître et de préserver ce savoir-faire artisanal, qui constitue le socle de la qualité des données médicales utilisées pour la recherche et la lutte contre le cancer.

La communauté des modérateur·rice·s de Stack Overflow

Avec votre aide, nous travaillons ensemble à construire une bibliothèque de réponses détaillées et de haute qualité à chaque question sur la programmation².
Stack Overflow

- 47 La seconde communauté que nous avons choisi d'étudier est celle des modérateur·rice·s bénévoles chargé·e·s de l'édition et de la validation des contenus produits sur la plateforme collaborative SO. Fonctionnant selon un modèle de questions-réponses, SO permet à chacun de poser une question sur un sujet (généralement technique), à laquelle tout membre de la communauté peut répondre (voir figure 3). Comme l'indique la promesse affichée sur la page d'accueil et reprise dans l'extrait ci-dessus, l'ensemble des membres œuvre à garantir la qualité du contenu informationnel.
- 48 Le travail de cette communauté de pratique repose sur un dispositif sociotechnique complexe — c'est-à-dire une infrastructure scripturale (Denis, *op. cit.*) — fait de règles, d'outils, d'un mécanisme de gratification et de rôles construits au fil du temps par la communauté. La communauté est ainsi régulée par un code de conduite³ reprenant les normes de forme et de fond auxquelles doivent se conformer tous les contenus publiés. Par ailleurs, comme l'illustre la figure 3, les utilisateur·rice·s de la plateforme ont à leur disposition des outils permettant de sanctionner une question ou une réponse par un vote ou un signalement. Ce système de vote est central au fonctionnement de la plateforme : il attribue des points positifs ou négatifs aux contenus et génère des points de réputation pour leurs auteur·rice·s, leur octroyant au fil du temps des outils et des rôles supplémentaires, comme le fait de pouvoir éditer des publications.

Figure . Exemples de fonctionnalités de modération sur Stack Overflow



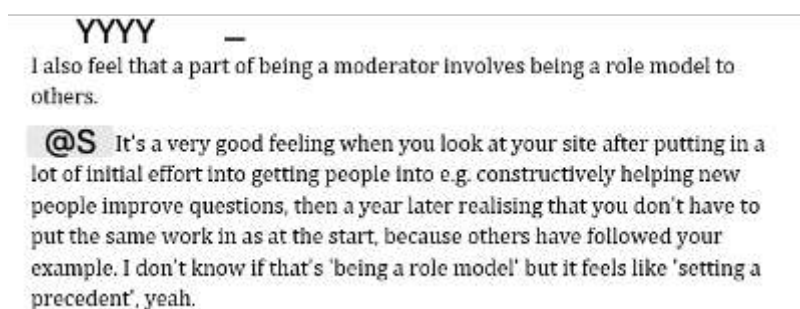
Légende : Stack Overflow, *Which JSON content type do I use?*, <https://stackoverflow.com/questions/477816/which-json-content-type-do-i-use>, page consultée le 11 avril 2025

- 49 La réflexivité joue un rôle central dans le travail de production et de modération. L'ensemble de la documentation nécessaire à ce travail ainsi que les discussions et décisions qui l'ont façonnée sont accessibles sur le site méta de la communauté⁴.

Récemment, à la suite de plusieurs différends avec l'entreprise qui héberge la plateforme, les modérateur·rice·s ont également mis en place un serveur Discord, où les membres discutent des règles et des outils, garantissant ainsi le bon fonctionnement de la communauté.

- 50 Ce qui frappe de prime abord lorsqu'on étudie cette communauté, c'est son organisation rigoureuse ainsi que les nombreux outils et procédures permettant à la fois de réguler l'afflux de contenus et le comportement des membres. Nous faisons ici l'hypothèse que ce dispositif sociotechnique complexe contribue en définitive à préserver le temps qualitatif de ces artisan·e·s bénévoles, qui peuvent ainsi se consacrer à l'accompagnement des utilisateur·rice·s (par l'entraide) et à l'amélioration du contenu informationnel (par l'édition de contenu).
- 51 Dans ce contexte, un mentorat par les pairs, assuré par les expert·e·s et les modérateur·rice·s, accompagne les personnes nouvellement membres. Comme l'illustre l'extrait ci-dessous (figure 4), issu d'une conversation sur le rôle des modérateur·rice·s, il existe une volonté affirmée d'agir de façon exemplaire pour les contributeur·rice·s ayant le moins d'expérience, au regard des standards de la communauté.

Figure 4. Extrait d'une conversation sur le rôle des modérateur·rice·s (Serveur Discord SO Moderators)⁵



- 52 L'attention portée aux contenus se manifeste à travers l'édition de contenu (les *handling flags* dans le jargon de la communauté). Cette exigence se traduit par une justification rigoureuse des décisions de modération, où chaque choix est argumenté et inscrit dans une logique de transparence et de fiabilité. L'édition des contenus témoigne ainsi d'un véritable travail sur les écrits, à l'image de celui des artisan·e·s façonnant la matière. Réalisé par les membres de la communauté, ce travail est au cœur de la préservation de la qualité et du soin apporté aux contenus présents sur la plateforme.

Figure . Question d'Artemis sur l'édition trop minutieuse d'un contenu⁶

What to do when a user targets your question history?

Asked 5 years, 3 months ago Modified 3 years, 9 months ago Viewed 674 times

▲ I'm not sure why, but recently, I had a single user (who only shares one tag with me) go through my question history, and make very questionable edits to posts that are both new (~1 month) and old (~8 months).

-5

▼ The changes are rather minute and unnecessary -

I am ~~very confused, and am looking for someone to help explain~~ that single line, and helping to map that one line into a much cleaner function. I am ~~seeking an explanation of~~ that single line, and helping to map that one line into a much cleaner function.

Why is this user scanning through posts of mine from months ago and making minute grammar changes? This makes me feel targeted.

Note - I should add, I am all in favor of keeping high quality QA on Stack (look through my comment history -- I often write about this to new contributors). If this is justified, that is fine, but I would like to better understand *why* that is a necessary addition without an explanation to the OP on how to more accurately improve their questions in the future.

discussion edits

Share Follow

edited Dec 5, 2019 at 19:10

asked Dec 5, 2019 at 16:30

 **artemis**
7,281 ● 9 ● 6

Figure . Réponse de Cody Gray défendant le travail d'édition des éditeur-ice-s⁷

▲ Yes, halfer does this. You aren't the first person to have noticed. :-)

40

▼ He's [a prolific editor](#), devoting a significant chunk of his time to improving posts on Stack Overflow. Anecdotally, I don't think I've ever seen him submit an edit that I didn't think was an improvement over the original.

halfer is, in my experience, particularly adept at removing "noise": trimming and rephrasing the author's words into something that is easier for others to read, without losing the original meaning. It looks like that's what he's done here, for your questions.

Rest assured that being "targeted" in this way (having your posts edited) is not a bad thing. It shouldn't be taken to assume that you've done anything wrong—just that someone saw a way they could improve the site and took the opportunity.

Stack Overflow is collaboratively edited; this is one of the site's core principles. In this regard, and others, we share a lot of DNA with Wikipedia. Our primary goal, in fact, is to build a library of high-quality answers to the long tail of programming questions. Improving the presentation of individual questions is an important mechanism for achieving this goal. You might think that the edits are "minor"—why would anyone care if the wording of a question is a bit rambling?—but remember that Q&A aren't just here to help the original asker. They are maintained over the long term, where they help hundreds or thousands of people who have the same question and find the answers on our site via search engines.

I would like to better understand *why* that is a necessary addition at the expense of the OP.

I would like to better understand why you think that edits are "at the expense of the OP". Edits *help* the OP, and they help everyone else.

Share Follow

answered Dec 5, 2019 at 17:52

 **Cody Gray** ♦ Mod
245k ● 84 ● 728 ● 770

13 that is exactly the type of response I was searching for. I appreciate the time you took to put together that response, and to attach and link some of the posts that you did. This is an excellent response. And, again for clarity, this was not an attack or knock -- *I just want to understand better so that I can one day make a bigger impact on the community.* Thanks! – [artemis](#) Dec 5, 2019 at 18:31

- 53 Dans la publication reproduite en figure 5, Artemis interroge le caractère parfois trop minutieux du travail d'un·e éditeur·rice et se demande si les efforts investis dans la

simple formulation d'une question présentent un réel intérêt. Dans la réponse qui s'ensuit (figure 6), Cody Gray illustre parfaitement la place centrale de l'édition dans la production de contenus de qualité. Il souligne ainsi que « ce que fait Halfer [l'éditeur-riche critiquée] consiste à enlever du bruit, à rendre les contenus plus faciles à lire pour autrui » et rappelle que l'édition collaborative est inscrite dans l'ADN de SO. C'est précisément ce travail qui, selon Cody Gray, permet de produire un contenu durable et de qualité.

- 54 SO répond ainsi pleinement aux critères d'une communauté de pratique : ses membres, tant les personnes les plus expérimentées que les nouvelles venues, partagent des objectifs communs, des méthodes, des valeurs, des règles et un vocabulaire spécifique. Les valeurs sous-jacentes de cette culture du « soin du contenu » s'ancrent dans l'éthique du *hacking*, qui promeut la création collaborative et décentralisée ainsi que la diffusion d'un savoir de qualité accessible à toutes et à tous (Turner, 2006).
- 55 En 2023, l'entreprise qui monétise les contenus produits par les bénévoles, en échange de la mise à disposition de l'infrastructure technique, a décidé d'interdire la suppression des contenus suspectés d'avoir été générés par l'IA. Cette décision a déclenché une grève des modérateur-riche-s de SO en réaction au manque de considération pour le travail accompli et parce que l'usage de l'IA générative les détourne de leur mission et de leur objectif.

Pour mes péchés, je suis le troisième plus grand éditeur de publications. L'édition semble désormais ancrée en moi, et je me demande si m'en défaire ne serait pas une épreuve. De plus, bien que nous, la communauté, voyions indéniablement la valeur de l'édition mutuelle, je ne suis pas certain que le retrait de ce travail préoccupe réellement l'entreprise⁸ (Halfer, notre traduction).

Je ne suis qu'un utilisateur « normal » qui utilise ses privilèges pour gérer le contenu au mieux de ses connaissances. Mais ces derniers jours, j'ai vraiment pris conscience de la situation et je me suis posé une question : Pourquoi est-ce que je continue à faire ça ? L'entreprise n'en a manifestement rien à faire, alors pourquoi devrais-je le faire ? C'est probablement la frustration qui parle, mais à ce stade, cela ne me dérangerait pas de laisser SE mourir dans une benne à ordures. À maintes reprises, les personnes qui se sont portées volontaires ici ont été déçues, alors pourquoi s'en préoccuper encore ? C'est pourquoi j'ai signé la lettre — Qbrute⁹ (notre traduction).

- 56 Le conflit autour des contenus générés par l'IA met en lumière plusieurs enjeux. Tout d'abord, comme illustré par le commentaire d'Halfer, cet épisode confirme que les modérateur-riche-s ont conscience de l'importance de leur travail dans le maintien d'un écosystème de savoir structuré, fiable et accessible. Dans ce contexte, les membres de la communauté établissent une nette distinction, en matière de qualité, entre les contenus produits et modérés par leurs soins et les contenus générés par l'IA. Par ailleurs, la grève apparaît comme une tentative pour faire prendre conscience à l'entreprise qui héberge la plateforme de l'importance de ce travail et des conséquences de sa décision quant à la qualité des contenus. Comme l'exprime Qbrute, la grève met ainsi en évidence le rôle des modérateur-riche-s et la nécessité de reconnaître leur contribution, mais aussi la nécessité de poursuivre le travail politique face à la pression grandissante de l'entreprise.
- 57 Enfin, la grève porte principalement sur la préservation du temps des modérateur-riche-s consacré aux pratiques artisanales des données. Or, en raison de la nouvelle politique de la plateforme, les modérateur-riche-s doivent désormais faire face à une surcharge de publications de faible qualité générées par l'IA, un phénomène qui détourne leur

énergie et leur expertise de leur mission première. En définitive, cette grève montre l'importance du rôle des modérateur·rice·s, non seulement aux yeux des utilisateur·rice·s, mais aussi de l'entreprise elle-même. La question de la qualité du contenu et de la reconnaissance du travail bénévole des artisan·e·s de la donnée se trouve ainsi au cœur des revendications.

Discussion et conclusion

- 58 En s'appuyant sur des données empiriques pour étayer notre proposition théorique, notre travail met en évidence plusieurs dimensions fondamentales de l'artisanat des données, qui seront discutées ici. Par ailleurs, nous proposons d'ouvrir la réflexion sur les tensions émergentes entre l'accumulation massive des données et l'exigence de qualité telle qu'on la trouve dans l'artisanat des données, en interrogeant la dimension politique de ce travail. Finalement et en guise de conclusion, nous reviendrons sur les limites et les perspectives qu'ouvre notre proposition théorique.

Retour sur la proposition d'artisanat des données

- 59 Bien que distincts dans leur finalité et leur organisation, les deux cas d'étude illustrent avec clarté que l'artisanat des données est une pratique sociale tout entière consacrée à la production de données de qualité. Les enquêteur·rice·s des RC, comme les modérateur·rice·s de SO, incarnent une forme d'artisanat de la donnée où l'expertise et le savoir-faire se construisent progressivement, à travers une interaction et une réflexivité constantes avec la matière première que représentent les données. L'artisanat des données, tel qu'on le trouve dans les deux cas d'étude, s'impose donc comme une pratique « à risque » (en opposition à la fabrication « certaine ») dans la mesure où ces pratiques artisanales opèrent dans un espace fait d'incertitudes, propice à l'expérimentation et à l'innovation, qui engage pleinement l'artisan·e dans son travail.
- 60 Ce souci constant de la qualité témoigne d'une attention minutieuse portée aux données, qui s'apparente à une forme de *care*, comme l'avaient déjà souligné plusieurs études (Stiefel, *op. cit.* ; Baker et Karasti, *op. cit.* ; Pinel, Prainsack et McKevitt, *op. cit.* ; Zakharova et Jarke, *op. cit.* ; Davies et Holmer, *op. cit.*). Dans cette perspective, le travail de soin s'incarne à travers des pratiques de maintenance et de réparation (Jackson, 2013 ; Denis et Pontille, 2022) des infrastructures de données et de leurs composants — en ce compris les données elles-mêmes (Baker et Karasti, *op. cit.*). Jackson met d'ailleurs l'accent sur cette filiation entre maintenance et soin porté aux choses en considérant que les pratiques qui en découlent réunissent « les mondes de l'action et du sens, reconnectant le travail nécessaire de maintenance avec les formes d'attachement qui, si souvent (mais de manière invisible, du moins pour les analystes), le soutiennent. Nous prenons soin parce que cela nous importe [*We care because we care*] » (*op. cit.*, p. 232, notre traduction).
- 61 Cette dimension s'avère d'autant plus prégnante que cette quête de qualité, couplée au soin des choses, revêt une portée éthique. En effet, les deux cas empiriques témoignent d'une conscience aiguë, chez les artisan·e·s, d'une forme d'éthique de la responsabilité dans la production et la gestion des données. Cette éthique de la responsabilité ne se limite pas à des considérations techniques, mais s'ancre dans un système de valeurs qui

guide leur travail au quotidien. Cette dimension est intimement liée à la finalité de leur activité, qu'il s'agisse de contribuer activement aux politiques de santé publique ou à l'utopie d'un Web ouvert, émancipateur et pourvoyeur de connaissances.

- 62 Cette responsabilité se manifeste de manière particulièrement marquée chez les enquêteur·rice·s et les représentant·e·s du CR, qui ne se contentent pas de produire des données fiables qui contribuent aux politiques de santé publique. Leur travail intègre aussi une réflexion critique sur la représentativité des minorités au sein des bases de données, afin d'éviter la reproduction d'inégalités systémiques. De manière analogue, dans la communauté SO, les artisan·e·s de la donnée s'investissent dans la démocratisation de l'apprentissage de la programmation. À travers le partage et la structuration de connaissances techniques sur la programmation, ces artisan·e·s bénévoles œuvrent pour un accès élargi aux savoirs, incarnant ainsi une vision émancipatrice de l'information et de la transmission du savoir.
- 63 Le concept d'artisanat des données permet toutefois de replacer cette éthique dans des dynamiques collectives. En effet, comme le montrent les deux cas d'étude, les artisan·e·s de la donnée évoluent également au sein de véritables communautés de pratique, rassemblant des membres aux niveaux d'expertise et d'expérience variables. Ces communautés forment le socle qui permet le développement des habiletés, elles favorisent les dynamiques d'apprentissage mutuel et de collaboration, où la production et la gestion des données sont des processus régulés et enrichis à travers l'échange avec d'autres membres. Ainsi, les artisan·e·s des données coconstruisent, partagent, valident et font circuler les connaissances, les méthodes, les valeurs selon une logique de coopération, garantissant non seulement la qualité des données produites, mais aussi la pérennisation (et, le cas échéant, l'évolution) des standards en vigueur dans la communauté.

Penser la dimension politique de l'artisanat des données

- 64 Dans les deux cas empiriques étudiés, la dimension politique de l'artisanat des données affleure à plusieurs reprises. Nous proposons de discuter cette dimension politique et son articulation avec les environnements datafiés à partir du concept de micropolitique. Développée par Deleuze et Guattari, la micropolitique repose sur l'idée que les centres de pouvoir sont multiples et que « tout est politique, mais toute politique est à la fois macropolitique et micropolitique » (1980, p. 260). Elle renvoie à la production de formes de pouvoir, de régulation et de gestion des relations sociales à une échelle locale, toujours en relation avec la macropolitique, qui s'exprime, quant à elle, à l'échelle des grandes structures institutionnelles et politiques.
- 65 La *datafication* (Mayer-Schönberger et Cukier, *op. cit.*), voire le dataïsme (van Dijck, 2014) relèvent de macropolitiques. Ils sont portés par des logiques institutionnelles, industrielles et gouvernementales qui privilégient l'accumulation massive, la standardisation et l'automatisation. L'artisanat des données peut alors être envisagé comme son actualisation à une échelle locale, où ces principes sont mis en œuvre, ajustés et contestés dans des pratiques concrètes de production et de gestion des données, orientées vers la qualité. Les artisan·e·s des données réintroduisent ainsi des formes de régulation échappant aux seules logiques industrielles, sans pour autant rejeter l'intérêt ou le potentiel des données et des technologies numériques dans leurs domaines respectifs. Leur travail consiste à négocier en permanence avec les

contraintes imposées par ces macropolitiques, tout en produisant, dans la pratique, des ajustements qui reflètent des préoccupations à la fois éthiques et pragmatiques.

- 66 Cette dimension politique s'exprime également dans la manière dont les artisan·e·s des données développent des savoirs situés et des pratiques de résistance discrètes face aux injonctions dominantes de la *datafication*. Dans un contexte où l'accumulation massive de données tend à être perçue comme une fin en soi, leur activité repose sur une exigence de qualité et de contextualisation, en rupture avec une vision extractiviste des données. Dans les deux cas étudiés, les artisan·e·s des données manifestent une vigilance particulière face aux risques que fait peser le déploiement de l'IA générative sur leurs pratiques et sur les valeurs qu'ils défendent. À l'instar du luddisme (Charitsis, Laamanen et Lehtiniemi, 2024), né au XIX^e siècle en réaction à la mécanisation du textile, l'IA est perçue comme une menace susceptible de conduire à la disparition des pratiques artisanales des données.
- 67 Cette tension est d'autant plus frappante que l'entraînement des IA génératives repose précisément sur des fondations solides : des données de qualité, sélectionnées et traitées avec soin. La dévalorisation apparente du travail de production et de gestion des données à l'ère de la *datafication* coïncide donc paradoxalement avec une dépendance accrue à ce travail pour garantir la fiabilité et la pertinence des résultats produits par les IA. Cette contradiction s'accroît encore avec l'essor des données synthétiques, c'est-à-dire artificiellement générées par l'IA, dont on sait désormais qu'elles alimentent des boucles autophages (ou « auto-consommatrices ») dégradant significativement la qualité des résultats produits par ces systèmes (Alemohammad *et al.*, 2023).

Limites et perspectives

- 68 Bien que certains traits artisanaux aient été relevés dans des recherches précédentes, notre analyse s'appuie seulement sur deux terrains empiriques et conserve donc, à ce stade, un caractère exploratoire. L'artisanat des données appelle dès lors à être étudié dans une pluralité de contextes empiriques. Notre proposition doit ainsi être entendue comme une invitation à approfondir l'investigation de ces pratiques par la recherche sur le travail artisanal des données. Au-delà de cette limitation empirique, plusieurs axes de réflexion méritent néanmoins d'être ouverts quant à la portée du concept.
- 69 Un premier enjeu concerne sa transférabilité. L'artisanat des données, tel qu'il se déploie dans les RC et sur SO, est-il susceptible de se manifester dans d'autres configurations de production ? La question se pose avec une acuité particulière dans les environnements caractérisés par une forte précarisation du travail, tels que ceux décrits par la littérature sur le *digital labor*. Cette dernière met en évidence des dynamiques d'exploitation, de standardisation et de fragmentation des tâches, qui semblent a priori incompatibles avec les caractéristiques de l'artisanat : autonomie relative, réflexivité, savoir-faire transmis et enrichi collectivement. Se pose dès lors la question de l'articulation entre ces deux registres : peut-on concevoir l'existence de zones de recoupement, où des micro-travailleurs élaborent, malgré tout, des pratiques artisanales (par exemple au travers de communautés informelles de soutien et d'entraide) ? Ou convient-il, au contraire, de penser l'artisanat comme un régime de travail radicalement distinct, réservé à des contextes institutionnels ou communautaires particuliers ?

- 70 Dans cette perspective, la finalité qui sous-tend l'artisanat des données apparaît également déterminante. Dans les RC, l'objectif de santé publique et la reconnaissance de l'expertise soutiennent une logique artisanale. Sur SO, la dimension communautaire, le bénévolat et l'idéologie open source favorisent la transmission et la valorisation des savoir-faire. Mais qu'en est-il dans les organisations privées, structurées par des impératifs commerciaux ? L'artisanat des données peut-il s'y épanouir alors que la logique de rentabilité tend à primer sur celle de la qualité ? Ou bien est-il condamné à être marginalisé, externalisé, voire subsumé par d'autres logiques organisationnelles ?
- 71 Enfin, la question de la reconnaissance sociale et institutionnelle se révèle décisive. L'artisanat des données demeure aujourd'hui largement invisible : il est faiblement valorisé dans les organisations, rarement reconnu dans les discours politiques et souvent occulté par la rhétorique de l'automatisation et les imaginaires de l'IA. Pourtant, comme nous l'avons montré, il constitue un maillon indispensable au fonctionnement des infrastructures numériques. Le saisir comme catégorie théorique permet non seulement d'en souligner la valeur propre, mais également d'ouvrir un horizon plus large lié à sa reconnaissance : comment rendre ce travail visible, lui accorder une reconnaissance sociale et institutionnelle, et mettre en place des dispositifs de protection et de soutien à cet artisanat à l'ère de la *datafication* ?
-

BIBLIOGRAPHIE

- ALEMOHAMMAD Sina, Josue CASCO-RODRIGUEZ, Lorenzo LUZI et Ahmed Imtiaz BARANIUK (2023), « Self-consuming generative models go mad », *arXiv preprint arXiv.2307.01850*, 4, p. 14, DOI : <https://doi.org/10.48550/arXiv.2307.01850>.
- BAKER Karen et Helena KARASTI (2018), « Data care and its politics: Designing for local collective data management as a neglected thing », *Actes du colloque 15th Participatory Design Conference: Full Papers-Volume 1*, p. 1-12.
- BAREIS Jascha et Christian KATZENBACH (2022), « Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics », *Science, Technology, & Human Values*, 47(5), p. 855-881.
- BARROWMAN Nick (2018), « Why data is never raw », *The New Atlantis*, 56, p. 129-135.
- BATES Jo, Yu-Wei LIN et Paula GOODALE (2016), « Data journeys: Capturing the socio-material constitution of data objects and flows », *Big Data & Society*, 3(2), DOI : <https://doi.org/10.1177/2053951716654502>.
- BELTRAME Tiziana et Ashveen PEERBAYE (2018), « Prendre soin des infrastructures. Introduction à la traduction de "L'ethnographie des infrastructures" de Susan Leigh Star », *Tracés. Revue de sciences humaines*, 35, p. 179-186.
- BENFELDT Olivia, John Stouby PERSSON et Sabine MADSEN (2020), « Data governance as a collective action problem », *Information Systems Frontiers*, 22, p. 299-313.

- BERNABEI Rina et Jacqueline POWER (2018), « Hybrid design: Combining craft and digital practice », *Craft Research*, 9(1), p. 119-134.
- BOCZKOWSKI Pablo J. et Kirsten A. FOOT (dir.), *Media technologies: Essays on communication, materiality, and society*, Cambridge, MIT Press, p. 221-239.
- BOSSEN Claus, Kathleen PINE, Federico CABITZA, Gunnar ELLINGSEN et Enrico Maria PIRAS (2019), « Data work in healthcare: An Introduction », *Health Informatics Journal*, 25(3), p. 465-474.
- CAI Li et Yangyong ZHU (2015), « The challenges of data quality and data quality assessment in the big data era », *Data Science Journal*, 14(2), p. 1-10.
- CARTER Daniel et Dan SHOLLER (2015), « Data science on the ground: Hype, criticism, and everyday work », *Journal of the Association for Information Science and Technology*, 67(10), p. 2309-2319.
- CASILLI Antonio (2015), « Digital labor : travail, technologies et conflictualités », dans Dominique CARDON et Antonio CASILLI (dir.), *Qu'est-ce que le Digital Labor ?*, Paris, Institut national de l'audiovisuel, p. 8-40.
- CHARITSIS Vassilis, Mikko LAAMANEN et Tuukka LEHTINIEMI (2024), « Towards algorithmic luddism: Class politics in data capitalism », *Information, Communication & Society*, 28, p. 1-18.
- CHOROSZEWICZ Marta (2022), « Emotional labour in the collaborative data practices of repurposing healthcare data and building data technologies », *Big Data & Society*, 9(1), DOI : <https://doi.org/10.1177/20539517221098413>.
- COOREN François (2015), « In medias res: Communication, existence, and materiality », *Communication Research and Practice*, 1(4), p. 307-321.
- CRAGIN Melissa et Kalpana SHANKAR (2006), « Scientific data collections and distributed collective practice », *Computer Supported Cooperative Work*, 15, p. 185-204.
- DAGIRAL Eric et Ashveen PEERBAYE (2022), « Les mains dans les bases de données. Connaître et faire reconnaître le travail invisible », *Revue d'anthropologie des connaissances*, 6(1), p. 191-216, DOI : <https://doi.org/10.3917/rac.015.0229>.
- DAVIES Sarah et Constantin HOLMER (2024), « Care, collaboration, and service in academic data work: Biocuration as “academia otherwise” », *Information, Communication & Society*, 27(4), p. 683-701.
- DE LA BELLACASA Maria Puig (2011), « Matters of care in technoscience: Assembling neglected things », *Social Studies of Science*, 41(1), p. 85-106.
- DELEUZE Gilles et Félix GUATTARI (1980), *Capitalisme et schizophrénie 2. Mille plateaux*, Paris, Éditions de Minuit.
- DENIS Jérôme (2018), *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines.
- DENIS Jérôme et David PONTILLE (2022), *Le soin des choses. Politiques de la maintenance*, Paris, La Découverte, coll. « Terrains philosophiques ».
- D'IGNAZIO Catherine et Lauren KLEIN (2023), *Data feminism*, Cambridge, MIT Press.
- DIMAS Geri Louise, Lauri GOLDKIND et Renata KONRAD (2023), « Big ideas, small data: Opportunities and challenges for data science and the social services sector », *Big Data & Society*, 10(1), p. 1-5.

- DOURISH Paul (2017), *The stuff of bits: An essay on the materialities of information*, Cambridge, MIT Press.
- FOSTER Jonathan, Julie MCLEOD, Jan NOLIN et Elke GREIFENEDER (2018), « Data work in context: Value, risks, and governance », *Journal of the Association for Information Science and Technology*, 69(12), p. 1414-1427.
- FYHN Håkon et Roger Andre SØRAA (2017), « Craftsmanship in the machine: Sustainability through new roles in building craft at the technologized building site », *Nordic Journal of Science and Technology Studies*, 5(2), p. 71-84.
- GITELMAN Lisa (2013), « *Raw Data* » is an oxymoron, Cambridge, MIT Press.
- GRAY Mary et Siddharth SURI (2019), *Ghost work. How to stop Silicon Valley from building a new global underclass*, Boston, Mariner Books.
- GRIMSHAW David (2017), « Crafting the Digital: Developing expression and materiality within digital design and manufacture », *The Design Journal*, 20(1), DOI : <https://doi.org/10.1080/14606925.2017.1352878>.
- HARPER Douglas (1987), *Working knowledge: Skill and community in a small shop*, Chicago, University of Chicago Press.
- HENKE Christopher (1999), « The mechanics of workplace order: Toward a sociology of repair », *Berkeley Journal of Sociology*, 44, p. 55-81.
- HOLMLID (dir.), *Nordes 2023: This space intentionally left blank*, 12-14 June, Linköping University, Norrköping, Sweden, DOI : <https://doi.org/10.21606/nordes.2023.47>.
- INGOLD Tim (2019), « Of work and words: Craft as a way of telling », *European Journal of Creative Practices in Cities and Landscapes*, 2(2), p. 5-17.
- IRANI Lilly (2015), « The cultural work of microwork », *New Media & Society*, 17(5), p. 720-739.
- JACKSON Steven (2014), « Rethinking Repair », dans Tarleton GILLEPSIE, Pablo J. BOCZKOWSKI, & Kirsten A. FOOT (Dir.), *Media technologies: Essays on communication, materiality, and society*, MIT Press Boston, p. 221-239.
- JARKE Juliane et Stefanie BÜCHNER (2024), « Who cares about data? Data care arrangements in everyday organisational practice », *Information, Communication & Society*, 27(4), p. 702-718.
- KLAMER Arjo (2012), « Crafting Culture: The importance of craftsmanship for the world of the arts and the economy at large », Actes du colloque *ACEI conference*, Kyoto, <http://www.klamer.nl/wp-content/uploads/2017/08/crafting.pdf>, page consultée le 9 octobre 2025.
- KOESTEN Laura, Kathleen GREGORY, Paul GROTH et Elena SIMPERL (2021), « Talking datasets – understanding data sensemaking behaviours », *International Journal of Human-Computer Studies*, 146, DOI : <https://doi.org/10.1016/j.ijhcs.2020.102562>.
- KOLKO Jon (2011), « Craftsmanship », *Interactions*, 18(6), p. 78-81.
- KUIJPERS Maikel H. G. (2019), « Material is the mother of innovation », dans Anna MIGNOSA et Priyatej KOTIPALLI (Dir.), *A Cultural Economic Analysis of Craft*, Cham, Springer International Publishing, p. 257-270.
- KUNNEMAN Harry (dir.) (2013), *Good work: The Ethics Of Craftsmanship*, Amsterdam, Uitgeverij SWP.

- LAMBOTTE François et Kévin CARILLON (2024), « Questionner les modes d'existence des traces numériques de nos échanges en organisation », *Approches théoriques en information-communication (ATIC)*, 8(1), p. 35-54.
- LATOUR Bruno (2004), « Why has critique run out of steam? From matters of fact to matters of concern », *Critical Inquiry*, 30(2), p. 225-248.
- LAUGIER Sandra et Pascale MOLINIER (2009), « Politiques du care », *Multitudes*, 37-38(2), p. 74-75.
- LAVE Jean et Etienne WENGER (1991), *Situated learning: Legitimate peripheral participation*, Cambridge, Cambridge University Press.
- LEONELLI Sabina (2016), « The philosophy of data », dans Luciano FLORIDI (dir.), *The Routledge handbook of philosophy of information*, New York, Routledge, p. 191-202.
- LEONELLI Sabina (2020), « Learning from Data Journeys », dans Sabina LEONELLI et Niccolò TEMPINI (dir.), *Data journeys in the sciences*, Berlin, Springer, p. 1-24.
- MACGILL Belinda (2019), « Craft, relational aesthetics and ethics of care », *Art/Research International: A Transdisciplinary Journal*, 4(1), p. 406-419.
- MARTIN Aryn, Natasha MYERS et Ana VISEU (2015), « The politics of care in technoscience », *Social Studies of Science*, 45(5), p. 625-641.
- MARTIN-SCHOLZ Anja, Anne MAYÈRE et François LAMBOTTE (2021), « Rendre compte du travail d'instauration et d'assemblage des données en santé : le cas de Registres des cancers en France », *Communication et organisation. Revue scientifique francophone en communication organisationnelle*, 59, p. 215-230.
- MASSANARI Adrienne (dir.), *Critical cyberculture studies*, New York, New York University Press, p. 255-269.
- MAYER-SCHÖNBERGER Viktor et Kenneth CUKIER (2013), *Big data: A revolution that will transform how we live, work, and think*, Boston, Houghton Mifflin Harcourt.
- MICELI Milagros, Martin SCHUESSLER et Tianling YANG (2020), « Between subjectivity and imposition: Power dynamics in data annotation for computer vision », *Actes du colloque ACM on Human-Computer Interaction*, 4(CSCW2), p. 1-25.
- MIGNOSA Anna et Priyatej KOTIPALLI (dir.), *A cultural economic analysis of craft*, Berlin, Springer International Publishing, p. 257-270.
- MULLER Michael, Ingrid LANGE, Dakuo WANG, David PIORKOWSKI, Jason TSAY, Vera LIAO, Casey DUGAN et Thomas ERICKSON (2019), « How data science workers work with data: Discovery, capture, curation, design, creation », *Actes du colloque CHI conference 2019 on human factors in computing systems*, New York, Association for Computing Machinery, p. 1-15.
- NACHTIGALL Troy Robert (2019), *Materializing data: Craftsmanship and technology for ultra-personalization*, thèse de doctorat sous la direction de Ron WAKKARY, Eindhoven, Technische Universiteit Eindhoven.
- PACHIDI Stella, Hans BERENDS, Samer FARAJ et Marleen HUYSMAN (2021), « Make way for the algorithms: Symbolic actions and change in a regime of knowing », *Organization Science*, 32(1), p. 18-41.
- PINEL Clémence, Barbara PRAINSACK et Christopher MCKEVITT (2020), « Caring for data: Value creation in a data-intensive research laboratory », *Social Studies of Science*, 50(2), p. 175-197.
- PYE Davide (1995), *The nature and art of workmanship*, Londres, Herbert Press.

- RENARD Damien, Salla-Maria LAAKSONEN, Sophie DEL FA, François LAMBOTTE et Jukka HUHTAMÄKI (2024), « Resistance(s) in platformed organizing », *40th EGOS Colloquium*, Milan, 4-6 juillet 2024, <https://dial.uclouvain.be/pr/boreal/object/boreal:292711>, page consultée le 12 septembre 2025.
- ROTHSCHILD Annabel, Amanda MENG, Carl DISALVO, Britney JOHNSON, Ben Rydal SHAPIRO et Betsy DISALVO (2022), « Interrogating data work as a community of practice », *Actes du colloque ACM on Human-Computer Interaction*, 6(CSCW2), p. 1-28.
- SCHUBERT Cornelius (2018), « Repair work as inquiry and improvisation: The curious case of medical practice », dans Philippe SORMANI, Alain BOVET and Ignaz STREBEL (dir.), *Repair work ethnographies: Revisiting breakdown, relocating materiality*, Singapore, Springer Nature Singapore, p. 31-60.
- SCHWALBE Michael (2010), « In search of craft », *Social Psychology Quarterly*, 73(2), p. 107-111.
- SENNETT Richard (2009), *The craftsman*, New Haven, Yale University Press.
- SHORTER Michael (2015), « The craft technologist », *Studies in Material Thinking*, 13, p. 1-19.
- STAR Susan Leigh (1999), « The ethnography of infrastructure », *American Behavioral Scientist*, 43(3), p. 377-391.
- STELMASZAK Marta (2022), « Inside a data science team: Data crafting in generating strategic value from analytics », *ECIS 2022 Research Papers*, 86, https://aisel.aisnet.org/ecis2022_rp/86, page consultée le 9 octobre 2025.
- STIEFEL Léa (2018), « Étudier le care en infrastructure. Les “petites mains” de la biobanque hospitalière », *Revue d’anthropologie des connaissances*, 12(3), p. 399-427.
- SUNDELIN Anders, Javier GONZALEZ-HUERTA, Krzysztof WNUK et Tony GORSCHKE (2021), « Towards an anatomy of software craftsmanship », *ACM Transactions on Software Engineering and Methodology*, 31(1), 6, p. 1-49.
- TREADAWAY Cathy (2007), « Digital crafting and crafting the digital », *The Design Journal*, 10(2), p. 35-48.
- TSAKNAKI Vasiliki et Anna VALLGÅRDA (2023), « Craft as a matter of care to inspire the design of computational things », *Actes du colloque Nordes 2023: This Space Intentionally Left Blank*, 12-14 juin, Linköping University, Norrköping, Sweden.
- TURNER Fred (2022) « Chapter 22. How digital technology found utopian ideology: Lessons from the first Hackers’ Conference », dans David SILVER et Adrienne MASSANARI (dir.), *Critical Cyberculture Studies*, New York, University Press, p. 255-269.
- VAN DIJCK José (2014), « Datafication, dataism and dataveillance: big data between scientific paradigm and ideology », *Surveillance and Society*, 12, p. 197-208.
- WAQUET Françoise (2022), *Dans les coulisses de la science : techniciens, petites mains et autres travailleurs invisibles*, Paris, CNRS Éditions.
- ZAKHAROVA Irina et Juliane JARKE (2024), « Care-ful data studies: Or, what do we see, when we look at datafied societies through the lens of care? », *Information, Communication & Society*, 27(4), p. 651-664.
- ZORAN Amit et Leah BUECHLEY (2013), « Hybrid reassemblage: An exploration of craft, digital fabrication and artifact uniqueness », *Leonardo*, 46(1), p. 4-10.

NOTES

1. <https://www.becquerel.fr/le-patient/les-registres-des-cancers/>, page consultée le 3 mars 2025.

2. <https://stackoverflow.com/tour>, page consultée le 10 mars 2025. Notre traduction.

3. Stack Overflow, *Code of Conduct*, <https://stackoverflow.com/conduct>, page consultée le 11 avril 2025.

4. Par exemple, Stack Overflow, *Top Questions*, <https://meta.stackoverflow.com/>, page consultée le 11 avril 2025.

5. « yyyy — 10/07/2020 15:06 - Je pense également qu'une partie du rôle de modérateur consiste à être un modèle pour les autres.

@S C'est très gratifiant de voir le résultat de tous les efforts que vous avez fournis au début pour inciter les gens à aider de manière constructive les nouveaux venus à améliorer leurs questions, puis de constater un an plus tard que vous n'avez plus besoin de faire autant d'efforts qu'au début, car d'autres ont suivi votre exemple. Je ne sais pas si c'est "être un modèle", mais cela ressemble à "créer un précédent", oui » (Notre traduction).

6. « Je ne sais pas trop pourquoi, mais récemment, un utilisateur (qui ne partage qu'un seul tag avec moi) a parcouru l'historique de mes questions et a apporté des modifications très discutables à des publications récentes (environ 1 mois) et anciennes (environ 8 mois).

Les modifications sont plutôt minimales et inutiles. Pourquoi cet utilisateur parcourt-il mes publications datant de plusieurs mois et apporte-t-il des modifications grammaticales minimales ? Je me sens visé.

Remarque — je tiens à préciser que je suis tout à fait favorable au maintien d'une qualité élevée sur Stack (consultez l'historique de mes commentaires, j'aborde souvent ce sujet avec les nouveaux contributeurs). Si cela se justifie, très bien, mais j'aimerais mieux comprendre pourquoi cette modification est nécessaire sans qu'aucune explication ne soit donnée à l'auteur du message initial sur la manière d'améliorer plus précisément ses questions à l'avenir. » Notre traduction.

7. « Oui, halfer fait cela. Vous n'êtes pas le premier à l'avoir remarqué. :-)

C'est un éditeur prolifique qui consacre une grande partie de son temps à améliorer les publications sur Stack Overflow. Pour l'anecdote, je ne pense pas l'avoir jamais vu soumettre une modification qui ne constituait pas une amélioration par rapport à l'original.

D'après mon expérience, halfer est particulièrement doué pour supprimer le "bruit" : il raccourcit et reformule les mots de l'auteur pour les rendre plus faciles à lire, sans perdre le sens original. Il semble que c'est ce qu'il a fait ici, pour vos questions. Soyez assuré qu'être "ciblé" de cette manière (voir vos publications modifiées) n'est pas une mauvaise chose. Il ne faut pas en conclure que vous avez fait quelque chose de mal, mais simplement que quelqu'un a vu un moyen d'améliorer le site et en a profité.

Stack Overflow est édité de manière collaborative ; c'est l'un des principes fondamentaux du site. À cet égard, et à d'autres, nous partageons beaucoup d'ADN avec Wikipédia. Notre objectif principal est en fait de constituer une bibliothèque de réponses de haute qualité à la longue liste de questions sur la programmation. L'amélioration de la présentation des questions individuelles est un mécanisme important pour atteindre cet objectif. Vous pensez peut-être que les modifications sont "mineures" (pourquoi quelqu'un se soucierait-il de la formulation d'une question un peu confuse ?), mais n'oubliez pas que les questions-réponses ne sont pas seulement là pour aider la personne qui a posé la question à l'origine. Ils sont conservés à long terme, où ils aident des centaines ou des milliers de personnes qui se posent la même question et trouvent les réponses sur notre site via les moteurs de recherche.

J'aimerais mieux comprendre pourquoi vous pensez que les modifications sont "au détriment de

l'auteur du message initial". Les modifications aident l'auteur du message initial, mais elles aident aussi tout le monde » (Notre traduction).

8. Stack Overflow, *How can we handle troublesome contributors during the moderation strike?*, <https://meta.stackoverflow.com/questions/425722/how-can-we-handle-troublesome-contributors-during-the-moderation-strike>, page consultée le 11 avril 2025.

9. Stack Overflow, *What has happened to lead moderators to consider striking?*, <https://meta.stackoverflow.com/questions/424979/>, page consultée le 11 avril 2025.

RÉSUMÉS

Cet article introduit et développe le concept d'artisanat des données pour qualifier un ensemble de pratiques, souvent invisibilisées, qui participent à la production des données à l'ère de la *datafication*. En contraste avec les discours dominants centrés sur la quantité et l'automatisation, l'artisanat des données met en lumière des formes de travail minutieux, situé et collectif, qui confèrent leur qualité et leur pertinence aux données. À partir de deux enquêtes ethnographiques, l'article montre que, dans certains contextes de production, le travail des données repose sur des savoir-faire, une éthique du soin (*care*) et des dynamiques communautaires comparables à celles que l'on trouve dans l'artisanat. Cette proposition théorique permet d'analyser le rôle des artisans des données et d'explorer les tensions entre massification des données, exigence de qualité et transformations liées à la *datafication* et à l'essor de l'intelligence artificielle (IA). L'artisanat des données apparaît ainsi comme une pratique à la fois technique, éthique et politique, qui vise à réaffirmer la dimension fondamentalement humaine du travail des données.

This article introduces and develops the concept of data craft to qualify a set of often invisible practices that contribute to data production in the age of datafication. In contrast to the prevailing narratives around quantity and automation, data craft highlights forms of meticulous, situated, and collective work that result in high-quality, relevant data. Drawing on two ethnographic surveys, the article shows that in certain production contexts, data work is based on know-how, ethics of care, and community dynamics comparable to those found in handicrafts. This theoretical position examines the role of data artisans and explores the tensions between data proliferation, expectations of quality, and transformations related to datafication and the development of artificial intelligence (AI). Data craft thus appears as a technical, ethical, and political practice, which aims to reaffirm the fundamentally human dimension of data work.

Este artículo incluye y desarrolla el concepto de artesanía de los datos para designar un conjunto de prácticas, a menudo invisibles, que intervienen en la producción de la información en la era de la *dataficación*. En contraste con los discursos dominantes centrados en la cantidad y automatización, la artesanía de datos pone en evidencia, formas de trabajo minucioso, localizado y colectivo, que confieren su calidad y pertenencia a los datos. Con base en dos encuestas etnográficas, el artículo resalta que en algunos contextos de producción, el trabajo de los datos se basa en el saber hacer, una ética de cuidado (*care*) y dinámicas comunitarias comparables a las que se encuentran en la artesanía. Esta propuesta teórica permite analizar el rol de los artesanos de los datos e identifica las tensiones entre la masificación de los datos, la exigencia de la calidad y las transformaciones relacionadas a la dataficación y expansión de la inteligencia artificial (IA).

Por lo tanto, la artesanía de datos aparece, de esta manera, como una práctica a la vez técnica, ética y política que tiene como objetivo reafirmar la dimensión fundamentalmente humana del trabajo de los datos.

INDEX

Mots-clés : artisanat des données, datafication, communauté de pratique, care, travail invisible

Palabras claves : artesanía de los datos, comunidad de práctica, care, dataficación, trabajo invisible

Keywords : data craft, community of practice, care, datafication, invisible work

AUTEURS

FRANÇOIS LAMBOTTE

François Lambotte est professeur en communication organisationnelle à l'UCLouvain (Belgique) et membre du laboratoire LASCO (Laboratoire d'analyse des systèmes de Communication des organisations). Ses recherches visent, entre autres, à mettre en lumière les enjeux humains et organisationnels qui sous-tendent les fabriques de données dans différents contextes scientifiques et techniques. Courriel : francois.lambotte@uclouvain.be

KEVIN CARILLON

Kevin Carillon est doctorant à l'UCLouvain (Belgique) en sciences de l'information et de la communication au sein du LASCO (Laboratoire d'analyse des systèmes de Communication des organisations). Ses recherches portent sur les processus de *datafication* et d'algorithmisation dans les médias de service public, avec une attention particulière portée à leurs implications organisationnelles. Courriel : kevin.carillon@uclouvain.be

Les études critiques sur les données

Généalogie et cartographie d'un champ

Florence Millerand, Mélanie Millette et Lucie Delias

- 1 Les études critiques sur les données, *critical data studies* en anglais, se structurent depuis les années 2010 autour du projet commun de déconstruire les discours présentant les données comme des éléments bruts et naturellement objectifs. Le champ a ainsi pris forme par le biais de recherches critiques qui montrent que les données sont des artefacts qui contribuent à mettre en forme la réalité dont elles témoignent, tout en étant formées par des structures sociales, culturelles et économiques (Gitelman, 2013 ; Hepp, Jarke et Kramp, 2022).
- 2 L'objectif de cet article est de proposer une cartographie des travaux en sciences humaines et sociales contribuant au champ émergent des études critiques sur les données, entendues au sens de *data*, c'est-à-dire les données produites en contexte numérique et exploitables notamment par des processus automatisés. La contribution de l'article est triple. Premièrement, il s'agit de mettre en valeur l'étendue de ce champ d'études, non seulement ses filiations, mais aussi ses spécificités par rapport à d'autres champs d'études critiques s'intéressant aux technologies numériques. Deuxièmement, nous souhaitons circonscrire les spécificités de ce champ d'études, notamment sur le plan épistémologique et de la constitution d'un objet propre. Troisièmement, il s'agit de présenter une lecture organisée de ces travaux en les regroupant autour d'axes thématiques. Ce faisant, l'article vise à dresser un bilan provisoire des recherches réalisées. Cet exercice de cartographie ne prétend pas à l'exhaustivité, même s'il s'efforce de présenter un portrait du champ. À ce titre, il porte une attention particulière aux recherches francophones en la matière, beaucoup moins nombreuses que celles en anglais. Disséminées dans des espaces disciplinaires variés, les publications francophones s'inscrivent rarement de manière explicite sous l'étiquette *critical data studies*, malgré leur apport manifeste — notamment autour des travaux fondateurs en sociologie de la quantification. Ainsi l'article permet-il de prendre en compte la contribution francophone au champ des études critiques sur les données, largement représenté par des travaux en anglais. Cette triple contribution de l'article vient corriger deux angles morts de la littérature : l'absence, à notre connaissance,

d'une cartographie des travaux critiques sur les données ainsi que la prise en compte des publications en français.

- 3 Sur le plan méthodologique, l'article repose sur une recension de la littérature qui a été réalisée en trois temps. Une première recension a été conduite à partir de recherches dans les principales bases de données en sciences humaines et sociales en 2021 à partir des mots-clés *critical data studies*, *datafication* et leurs équivalents en français : *études* ou *approches critiques sur les données*. Par la suite, cette recension a été complétée au fur et à mesure de la découverte de nouvelles références pertinentes, pour atteindre un total d'environ 250 références parues entre 2010¹ et 2024. Une seconde recension a été réalisée en janvier 2025 pour mettre à jour la liste des références, avec deux objectifs : ajouter les travaux parus depuis 2021 en anglais dans un souci de complétion et repérer de façon plus fine les travaux en français parus entre 2015 et 2025. Pour ces derniers, la liste des mots-clés a été élargie pour inclure, entre autres, les travaux critiques sur la quantification et les biais algorithmiques². Troisièmement, une recension systématique et une analyse bibliométrique ont été menées dans la base de données Scopus sur l'ensemble des références contenant l'expression « *critical data studies* » (en anglais) dans leur titre, leur résumé ou leurs mots-clés. Cette recherche a permis d'obtenir 297 références parues entre 2015 et 2025, dont la majorité recoupait celles déjà référencées, tout en permettant le repérage de nouveaux travaux³. Une analyse bibliométrique a confirmé certains de nos constats sur l'organisation du champ, notamment l'identification des articles fondateurs, l'évolution du nombre de publications dans le temps et les champs disciplinaires les plus représentés à partir des revues les plus importantes⁴. L'ensemble des références recensées ont été triées, analysées et classées manuellement afin de dégager les grands thèmes et problématiques qui organisent le champ des études critiques sur les données.
- 4 Malgré ce travail méthodologique rigoureux et la réalisation de trois recensions successives de la littérature, l'article comporte quelques biais. Par exemple, les bases de données comme Scopus traitent beaucoup plus efficacement les articles en anglais. Ensuite, le champ s'est d'abord doté d'une étiquette en anglais, ce qui facilite le repérage dans cette langue puisque les publications subséquentes s'identifient comme telles. L'article reflète donc non seulement l'asymétrie anglais/français présente dans le champ, mais aussi celle des outils bibliométriques souvent moins précis avec les langues autres que l'anglais.
- 5 L'article s'organise en deux parties. La première présente le contexte d'émergence des études critiques sur les données en tant que champ de recherche, dans les espaces scientifiques anglophones et francophones, puis précise la perspective épistémologique du champ de même que les spécificités de son objet de recherche. La deuxième partie organise une sélection des travaux en identifiant deux grands axes de problématiques : l'étude des conséquences de la mise en données généralisée sur la vie sociale et l'étude des solutions de remplacement à la *datafication*.

L'émergence du champ des études critiques sur les données

- 6 Les études critiques sur les données constituent un champ interdisciplinaire relativement jeune, puisqu'il commence à se structurer à partir des années 2010 en lien avec l'accélération de la production et de l'exploitation des données numériques, dans

la foulée de l'adoption massive des applications mobiles et de la plateformes des médias sociaux associés au *big data* (Kitchin, 2022). Dans le milieu scientifique, le terme *datafication* se développe à la même époque afin de désigner les conséquences de la prédominance croissante des données dans l'ensemble des domaines de la vie sociale (Mayer-Schönberger et Cukier, 2013 ; Lycett, 2013 ; Van Dijck, 2014). Traduit en français par l'expression « mise en données », la *datafication* renvoie, d'un côté, aux « processus par lesquels les sujets, les objets et les pratiques sociales sont transformés en données numériques⁵ » (Southerton, 2022, p. 358, notre traduction) et, de l'autre, à une logique économique et politique qui envisage le monde comme une source de données à exploiter, à des fins diverses (commerciales, scientifiques, etc.) (Crawford, 2021 ; Mattelart, 2019). Dans ce contexte, les recherches se penchent sur les implications de cette profusion de données, produites en direct, non seulement à l'échelle des organisations et des États, mais aussi par les individus dans leur vie quotidienne (Lupton, 2015).

- 7 La dénomination *critical data studies* apparaît en 2014 dans un article publié dans le magazine scientifique en ligne *Society & Space*, signé par Dalton et Thatcher, deux chercheurs en géographie. Dans ce court texte se voulant provocateur, les auteurs proposent une critique du régime de données instauré par le *big data*. Ils évoquent notamment l'importance d'historiciser le développement des données massives pour en comprendre les effets sur la société, la nécessaire déconstruction de l'apparente objectivité des données, le danger de tomber dans des formes de déterminisme technologique et l'urgence d'imaginer des régimes alternatifs et des « contre-données ». Avant eux, les articles de boyd et Crawford avaient posé les bases d'un programme de recherche pour les études critiques sur les données, en formulant une critique de l'épistémologie du *big data*, de la rhétorique utopique et dystopique l'entourant ainsi qu'en soulevant la nécessité de questionner le pouvoir des données à (re)produire la réalité (boyd et Crawford, 2012 ; Crawford, Gray et Miltner, 2014).
- 8 Dans les années suivantes, un programme de recherche concernant les études critiques sur les données prend forme à travers une série de publications phares, dont un article de Kitchin, chercheur en géographie, et Lauriault, chercheuse en communication, qui reprend les idées de Dalton et Thatcher (2014) en appelant à mener un travail conceptuel arrimé à la conduite d'études empiriques pour développer le champ (Kitchin et Lauriault, 2014). Markham, Tiidenberg et Herman codirigent en 2018 un numéro spécial de la revue influente *Social Media + Society* sur le thème *Doing ethics in the era of big data research*, regroupant des textes qui interrogent les perspectives éthiques du *big data*, prolongeant ainsi les interrogations et réflexions critiques soulevées par boyd et Crawford dès 2012 (Luka et Millette, 2018). En 2016, un numéro spécial de la revue *Big Data & Society* est consacré pour la première fois aux *critical data studies* (Iliadis et Russo, 2016), avec, entre autres, un entretien entre Dalton, Thatcher et Taylor, chercheuse en droit des technologies, qui discute des grands enjeux de recherche à examiner, notamment la nature profondément spatiale des données, la variabilité historique des processus de production et d'accumulation des données ainsi que le développement de régimes de données pilotés par l'industrie de la *big tech* (Dalton, Taylor et Thatcher, 2016). Ce numéro spécial de *Big Data & Society* marque un jalon important : les *critical data studies* constituent désormais un champ de recherche à part entière. D'ailleurs, les publications se multiplient par la suite, passant d'une quinzaine

annuellement en 2016 à une trentaine en 2020, puis une soixantaine en 2022 (voir l'annexe 1).

- 9 À partir des années 2020, plusieurs ouvrages collectifs sont publiés, dont *New perspectives in critical data studies: The ambivalences of data power* (Hepp, Jarke et Kramp, *op. cit.*), qui propose un premier bilan de l'état des recherches. L'ouvrage réitère l'appel formulé dans les textes fondateurs : aborder de manière critique le pouvoir des données pour en comprendre les ambivalences. Les contributions rassemblées dans le collectif invitent à réfléchir, d'un côté, aux nouvelles occasions d'action offertes par les données à l'échelle de la société, par exemple lorsque des données de consommation d'énergie permettent de soutenir des processus plus durables, et, de l'autre, aux risques qu'elles suscitent en matière de surveillance, de vie privée et de reproduction des inégalités sociales. Depuis 2025, signe de la maturité du champ, les études critiques sur les données disposent de leur propre dictionnaire, *Critical data studies. An A to Z guide to concepts and methods* (Kitchin, 2025), regroupant plus de 400 termes et concepts clés allant des algorithmes à la sousveillance en passant par la *datafication*. Signalons également le *SAGE handbook of data and society* publié en 2025 (Venturini *et al.*, 2025) et le *De Gruyter handbook of critical data studies* (Mejias, McNealy et Miceli, à paraître en 2026).
- 10 Cette production scientifique s'est développée dans des espaces de discussion qui ont largement contribué à structurer le champ tel qu'il se présente aujourd'hui, principalement en anglais. La revue *Big Data & Society*, fondée en 2014 par la sociologue Evelyn Ruppert, a publié le plus grand nombre d'articles — et de loin — s'inscrivant explicitement dans le champ (voir l'annexe 2). Des sociétés savantes (comme la Society for Social Studies of Science — 4S), des instituts de recherche (comme le Data & Society Research Institute, fondé par danah boyd), des initiatives militantes (comme *Data For Black Lives*) et des conférences ont également joué un rôle structurant. La conférence *Data Power* en particulier, dont la première édition s'est tenue en 2015, a réuni les chercheurs et chercheuses du domaine de façon régulière. Désormais à sa cinquième édition, elle a donné lieu à plusieurs publications ; citons, à titre d'exemples, les numéros spéciaux « Data power in material contexts » publié dans la revue *Television and New Media* (Kennedy et Bates, 2017) et « Data power » paru dans la *Online Information Review* (Gerrard et Bates, 2019) ainsi que dans le *Canadian Journal of Communication* (Lauriault et Lim, 2019) et l'ouvrage collectif de Hepp, Jarke et Kramp (*op. cit.*) issu de la conférence de 2022.
- 11 Sur le plan des filiations théoriques, les travaux en études critiques des données se situent, du côté anglophone, au croisement de différentes approches issues de l'histoire des sciences, des *sciences & technology studies* (STS), de la géographie, de l'anthropologie, des études des archives, de la communication, des études médiatiques et des *Internet studies*, plus spécifiquement les sous-champs de l'éthique d'Internet et des études sur les plateformes. Autrement dit, les études critiques des données émergent au croisement de champs et sous-champs disciplinaires en sciences humaines et sociales qui se démarquent par le regard critique qu'ils portent sur la production de connaissances et la technologie. L'influence des STS et de l'histoire de la science se fait sentir par la publication de certains ouvrages clés sur l'histoire des données qui ont pavé la voie aux études critiques, notamment « Raw data » is an oxymoron, de Lisa Gitelman (*op. cit.*). Des travaux en *Internet studies* ont également contribué à formuler une critique du travail de cadrage de la réalité qui s'effectue à travers l'activité scientifique de production de données (Markham, 2013). La contribution de géographes

critiques a donné lieu à des textes fondateurs inspirés, entre autres, par un regard réflexif sur la responsabilité de leur domaine dans l'essor des données massives, par exemple à travers la généralisation des systèmes de géolocalisation (systèmes d'information géographique) (Dalton et Thatcher, 2014) et le développement d'assemblages sociotechniques complexes, comme les villes connectées (Kitchin, 2014a), sources importantes de production et de consommation de données.

- 12 En français, la production de travaux relevant des études critiques sur les données est à la fois moindre et plus dispersée. Il faut dire que les recherches sur le sujet ne revendiquent pas forcément leur inscription sous la dénomination des *critical data studies* (ou son équivalent en français), un champ qui reste issu de la recherche anglophone. Sur le plan des publications, citons le numéro « Production des données. Production de la société » publié par la revue *Les Enjeux* (Bullich et Clavier, 2018) ainsi que le numéro « Produire les données, à la frontière de plusieurs mondes sociaux » de la revue *Réseaux* (Flichy, 2021). *Réseaux* publiera d'ailleurs d'autres numéros touchant à des problématiques connexes, comme les villes intelligentes (Bernardin et Jeannot, 2019) et la sociologie des bases de données (Flichy et Parasie, 2013). Citons également le numéro récent « Big Data : usages et résistances », paru dans la revue *Quaderni* (Theviot et Bellon, 2024).
- 13 Sur le plan des ancrages théoriques, plusieurs travaux critiques sur les données en français viennent prolonger la frange critique de la sociologie de la quantification (Henneguelle et Jatteau, 2021). Rappelons que la question de savoir ce qu'est une « donnée » a alimenté un ensemble de recherches sur les données et les chiffres bien avant l'ère numérique (Mattelart, *op. cit.*). Par exemple, la sociologie de la quantification et de la statistique, à travers les travaux fondateurs de Desrosières (2000) et, plus récemment, de Martin (2020), a mis en lumière la « valeur sociale » accordée aux chiffres dans les sociétés occidentales, que les travaux sur les quantifications de soi réactualisent en partie (Dagiral *et al.*, 2019). Le parallèle est intéressant à faire entre la prolifération actuelle des données et les raisons évoquées par Martin (*ibid.*) pour expliquer l'expansion de la quantification : quantifier pour connaître, pour répondre à une quête de justice, pour se coordonner, pour établir la confiance, pour exercer le pouvoir, etc. Tous ces motifs peuvent s'appliquer aux régimes de données contemporains (Mattelart, *op. cit.*).
- 14 D'autres travaux en français s'inscrivant dans les STS se sont penchés sur l'histoire et la nature des données scientifiques, en les qualifiant d'« obtenues » plutôt que de « données », selon la formule de Latour (2007). L'ouvrage de Denis (2018), *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, constitue un bon exemple de prolongement de ces questionnements critiques sur les données numériques. Dans une autre perspective, l'ouvrage collectif *Gouverner par les données ? Pour une sociologie politique du numérique*, dirigé par Theviot (2023), propose de déconstruire les mythes autour du *big data* en interrogeant la manière dont les données sont utilisées ainsi que les modalités de leur influence sur les modes de gouvernance et les prises de décision dans les organisations étatiques, marchandes et scientifiques. La critique des données massives alimente également plusieurs travaux en communication, notamment sur l'illusion de neutralité des données, en mettant en évidence la façon dont cette apparence de neutralité sert des intérêts économiques et de pouvoir (Crémier, Bonenfant et Lafrance, 2019 ; Ménard *et al.*, 2018). Des travaux récents autour des « publics de données » s'inscrivent dans ces prolongements critiques

en déplaçant l'attention des données vers les individus, les collectifs et les organisations représentés dans les données ou à qui elles sont destinées (Millerand *et al.*, 2025).

- 15 Ce bref historique du champ permet d'identifier les traditions disciplinaires dans lesquelles s'ancrent les études critiques des données. Si les fondateurs viennent de la géographie et des *Internet studies*, plusieurs travaux se situent aujourd'hui à la croisée de la communication et des études médiatiques (sciences de l'information et de la communication en France), de la sociologie, de l'anthropologie et des STS. Soulignons à ce sujet que les revues qui publient le plus de travaux (en anglais) s'inscrivant dans les *critical data studies* sont en majorité des revues interdisciplinaires comme *Big Data & Society* et *Internet Policy Review* ainsi que des revues associées à la communication comme *Information, Communication & Society*, *New Media & Society* et *International Journal of Communication* (voir l'annexe 2). Même si tous les travaux en français ne font pas référence de manière explicite au champ des *critical data studies*, certains citent des publications phares du champ ou épousent les mêmes présupposés épistémologiques, à savoir un regard critique sur les données et sur les mutations sociales, économiques et politiques qui les accompagnent. À ce titre, certains auteurs et autrices, y compris anglophones, s'inscrivent aussi ou plus explicitement dans des champs connexes marqués par une approche critique des technologies numériques, comme les *infrastructure studies* (Bowker *et al.*, 2010 ; Knox et Gambino, 2023), les *critical algorithm studies* (Gillespie, 2014, 2015 ; Beer, 2017), les *platforms studies* (Gillespie, 2018 ; Burgess, 2021) ou plus récemment les *critical AI studies* (Lingren, 2023 ; Crawford, *op. cit.*). À l'instar des *critical data studies* qui placent les *data* au cœur des questionnements de recherche, ces champs visent à attirer l'attention de la communauté de recherche sur des objets et dispositifs spécifiques (les infrastructures, les algorithmes, les plateformes, l'IA) et à proposer des programmes de recherche transversaux aux sciences humaines et sociales. Face à cette prolifération et à cet éclatement apparent des champs d'études — qui se multiplient depuis les années 2000 (Silver, 2004), les études critiques sur les données proposent ainsi une approche transversale des enjeux sociaux, politiques et économiques liés aux technologies numériques, dont l'exploitation des données est à la base du fonctionnement.

Les données comme objet de recherche

- 16 Sur le plan épistémologique, nous identifions deux spécificités des études critiques sur les données — tant dans la recherche anglophone que francophone. D'abord, une épistémologie intrinsèquement sociotechnique, revendiquée dans les travaux fondateurs (Dalton, Taylor et Thatcher, *op. cit.*) comme dans les plus récents (Kitchin, 2025), où l'on appréhende la constitution des données comme résultant à la fois d'une pratique et d'un travail réalisé par des humains portant leur subjectivité et par la mobilisation de différents dispositifs de collecte et de traitement des données. L'interinfluence de ces deux facettes s'avère constitutive de la perspective épistémologique du champ, alors que les dispositifs de collecte, de gestion et de visualisation des données, de même que les décisions humaines associées, sont souvent naturalisés ou invisibilisés, et cela, dans une foule de domaines (marketing, sécurité, finance, santé, etc.) où les données jouent un rôle majeur et sont comprises a priori comme étant le témoin neutre d'une réalité. Ensuite, les études sur les données

s'inscrivent, comme leur nom l'indique, dans une épistémologie critique qui se démarque par une perspective sur le monde centrée sur les rapports de pouvoir, leur reproduction et leurs répercussions, dans une visée de justice sociale (boyd et Crawford, *op. cit.* ; Luka et Millette, *op. cit.* ; Fuchs, 2019). Cette spécificité épistémologique s'inscrit dans le prolongement de certaines perspectives des travaux fondateurs, comme l'éthique, les études trans, féministes et postcoloniales (Markham, Tidenberg et Herman, 2018 ; Hoffman, 2020 ; D'Ignazio et Klein, 2020 ; Marcetic et Nolin, 2022 ; Mejias et Couldry, 2024), mais aussi en rupture avec certaines approches des données qui s'avèrent peu critiques ou acritiques, comme les perspectives post-positivistes ou issues du marketing. Cette spécificité épistémologique a également contribué à cadrer l'objet, les données, de manière critique en les abordant comme la manifestation de « formes de pouvoir » (Iliadis et Russo, *op. cit.*, paragr. 1, notre traduction). Si l'on s'attarde aux différents apports qui inscrivent l'épistémologie critique au fondement du champ, le féminisme intersectionnel contribue notamment à la critique de la reproduction de biais sexistes et racistes, mais aussi capacitistes et âgistes, dans les jeux de données et les systèmes qui les mobilisent (Noble, 2018 ; D'Ignazio et Klein, *op. cit.*). Les perspectives dé-coloniales et postcoloniales abordent la *datafication* de la société par une critique de l'extraction de données à des fins lucratives, en établissant une continuité entre le colonialisme et le capitalisme (Taffel, 2021). Tout comme le colonialisme s'approprie des territoires et des ressources à des fins de profit, le colonialisme par les données « normalise l'exploitation des êtres humains par le biais de données » (Couldry et Mejias, 2018, p. 336). Plus largement, les perspectives issues de l'économie politique abordent la place centrale des données dans la société par l'examen critique des changements opérés dans les modes de gouvernance et de la gouvernamentalité, de la prise de décision basée sur les données dans les bureaucraties aux processus de surveillance généralisée (Ouellet *et al.*, 2015 ; Mattelart, *op. cit.* ; Fuchs, *op. cit.*). Plus récemment, l'inscription des données dans des rapports de pouvoir préexistants, relus au prisme de la *datafication*, a donné à voir un « capitalisme de surveillance » (Zuboff, 2019), une nouvelle forme de « colonialisme » par les données (Mejias et Couldry, *op. cit.*) et le travail des algorithmes entraînés sur celles-ci, sources de biais potentiellement préjudiciables (Noble, *op. cit.* ; Broussard, 2023). Il s'agit, à travers ces publications, de déconstruire les fondements idéologiques de la *datafication*, en particulier la croyance répandue dans le *dataism* (qui suppose que le monde peut être compris à travers les data) en mettant au jour l'imbrication du politique (les gouvernements), de l'économique (les plateformes) et du scientifique (les chercheurs et chercheuses) au cœur de cette idéologie (Van Dijck, *op. cit.*).

- 17 Le caractère contextuel des données et leur nature éminemment politique (Bigo *et al.*, 2019) sont deux présupposés qui fondent l'objet de recherche privilégié du champ et qui teintent par le fait même son épistémologie critique. Les premiers travaux ont cherché à déconstruire activement les discours neutralisants sur la révolution du *big data*, qui présentaient les données comme une nouvelle matière première brute, neutre et exploitable, et leur traitement automatisé comme une source de connaissance inédite aux logiques des comportements humains (Kitchin et McArdle, 2016). Le caractère situé des régimes de production de données, soumis à diverses logiques de pouvoir (boyd et Crawford, *op. cit.*), et le caractère construit de la soi-disant « donnée brute » ont été largement documentés (Gitelman, *op. cit.* ; Denis, *op. cit.*). Ainsi,
les [*critical data studies*] considèrent que les données sont contingentes et relationnelles par nature ; leur génération, leur traitement et leur utilisation

s'inscrivent dans un contexte (conceptuel, technique, social, économique) et sont profondément imprégnés de politique (personnelle, institutionnelle, culturelle)⁶ (Kitchin, 2025, p. 41, notre traduction).

- 18 Cette manière de problématiser l'objet de recherche du champ implique « [...] qu'une donnée constitue toujours un artefact de médiation, c'est-à-dire une construction scientifique qui témoigne effectivement d'une partie de la réalité, qui la représente, sans être exactement cette réalité » (Millette, 2023, p. 15).
- 19 Dans cette perspective, les données, tout en étant un moyen de connaître et de comprendre le monde, constituent également des « matériaux sociotechniques contingents et changeants, essentiels à la production de connaissances, à la création de valeur et à la gestion de la société » qu'il s'agit de dénaturer et de repolitiser (Kitchin, 2025, p. 41).
- 20 Mais quelles sont les conséquences méthodologiques de l'adoption d'une telle posture épistémologique sociotechnique et critique ? Hepp, Jarke et Kramp (*op. cit.*) soulignent à ce sujet que la réflexion méthodologique sur l'examen critique du pouvoir des données reste centrale depuis l'émergence du champ. Il s'agit alors, pour les chercheurs et chercheuses, d'étudier les pratiques autour des données et de dégager les épistémologies qui les sous-tendent, et, dans la foulée des appels des textes fondateurs, d'entreprendre un « travail ethnographique et discursif » pour décrire non seulement les données, les cultures qui les entourent et les systèmes algorithmiques (Dalton, Taylor et Thatcher, *op. cit.*, p. 7), mais aussi les contextes de production des données, y compris la question politique de déterminer ce qui mérite d'être mis en données et ce qui ne le mérite pas (D'Ignazio et Klein, *op. cit.*). L'entretien d'un certain scepticisme à l'égard des données massives et de l'emploi des seules méthodes computationnelles marque le champ depuis ses débuts. Kitchin et McArdle (*op. cit.*) y ont directement contribué par une méta-analyse de plus d'une vingtaine de recherches basées sur l'analyse de données qualifiées de massives, en montrant qu'elles partageaient, finalement, très peu de caractéristiques propres au *big data*, tant sur le plan méthodologique que des jeux de données. Cette démarche ayant démontré que ce qui est considéré comme du *big data* dans une discipline ne l'est pas du tout dans une autre, les chercheurs et chercheuses en études critiques des données ont privilégié une diversité de méthodes, notamment qualitatives, et la pratique de la triangulation méthodologique (Hepp, Jarke et Kramp, *op. cit.* ; Latzko-Toth, Bonneau et Millette, 2017). Ces perspectives méthodologiques sont cohérentes avec l'épistémologie critique du champ, alors que de petits jeux de données denses (*thick data*) permettent, par exemple, de mieux cerner la mise en données d'un phénomène complexe et de le replacer dans les enjeux de pouvoir qui le traversent (Latzko-Toth, Bonneau et Millette, 2020 ; Christin, 2020 ; Millette, *op. cit.*).

Les thèmes de recherche structurant le champ

- 21 À partir de nos recensions des travaux relevant des études critiques sur les données, nous présentons dans cette section les thèmes de recherche principaux de ce champ d'études, organisés autour de deux axes structurants. Le premier porte sur l'analyse des conséquences de la *datafication*, tandis que le deuxième s'intéresse aux solutions de remplacement à ce phénomène.

L'étude des conséquences de la mise en données de la société

- 22 L'une des préoccupations centrales des travaux critiques sur les données réside dans la mise au jour et l'analyse des conséquences de la mise en données généralisée sur la vie sociale, envisagée dans son acception la plus large. Dans cette section, il s'agira de proposer une synthèse des recherches, à la dimension souvent empirique, qui s'intéressent à la manière dont l'introduction des données et leur traitement automatisé dans différents contextes affectent « leurs "sujets", c'est-à-dire [les entités], les personnes et les activités concernées par cette quantification » (Christin, *op. cit.*, p. 1116). Nous reviendrons d'abord sur les travaux portant sur les conséquences économiques, politiques et sociales de la *datafication* qui ont émergé à mesure que le traitement automatique des données s'est imposé parallèlement dans la société et dans la recherche, puis sur ceux abordant ses conséquences sur les pratiques et secteurs professionnels.

Les conséquences économiques, politiques et sociales

- 23 À partir du début des années 2010, les recherches en *critical data studies*, en lien avec l'émergence des discours autour de l'explosion du *big data*, commencent à s'intéresser aux conséquences de la *datafication* sur l'économie et sur les mécanismes de production de valeur (Lycett, *op. cit.*) ainsi qu'aux liens d'interdépendance entre le capitalisme, le *big data* et la marchandisation des corps et des comportements individuels. Plusieurs auteurs théorisent le renouvellement du capitalisme « traditionnel » face à un système où les traces laissées en ligne par les internautes et celles générées par leurs comportements deviennent une nouvelle matière première à exploiter. C'est le cas du « capitalisme de surveillance » (Zuboff, *op. cit.*) qui met l'accent sur l'orientation des comportements individuels à partir des données personnelles, du « techno-féodalisme » (Durand, 2021) qui décrit la manière dont les entreprises du Web vivent de la rente du travail gratuit des utilisateurs de leurs services ou encore du « colonialisme de données » (Meijas et Couldry, *op. cit.*) qui met en avant le prolongement des logiques extractivistes des empires coloniaux, désormais incarnés par les géants du numérique. Plus précisément, les recherches explorent la manière dont les données sont constituées en actifs économiques pour les entreprises (Birch, 2023) ou s'intéressent à des acteurs spécifiques comme les courtiers de données (*data brokers*) (Crain, 2018).
- 24 Les travaux en *critical data studies* s'interrogent également sur les transformations apportées par les données dans le champ politique. Dès la première moitié des années 2010, des recherches commencent à analyser de manière critique le mouvement des données ouvertes (*open data*) et de l'application de ses principes par les gouvernements à différentes échelles (États, municipalités, institutions internationales). La philosophie de ce mouvement repose sur l'idée que l'ouverture du plus grand nombre possible de jeux de données produits par les institutions publiques aurait des effets bénéfiques dans plusieurs domaines : renouvellement de la démocratie grâce à une plus grande transparence de l'action publique et encapacitation citoyenne (Badouard, 2017 ; Baack, 2015), augmentation de l'efficacité des administrations (Chevallier, 2018 ; Purwanto, Zuiderwijk et Janssen, 2020) et stimulation de l'innovation technologique ainsi que de l'activité entrepreneuriale (Cardulo et Kitchin, 2019).
- 25 Certaines recherches invitent à nuancer les discours optimistes, voire naïfs, concernant l'effet de l'ouverture des données, celle-ci n'étant pas automatiquement associée à une

augmentation de la puissance d'agir des citoyens (Goëta et Mabi, 2014), et pourrait au contraire générer une forme de démocratie fondée sur une action citoyenne se limitant à la surveillance d'indicateurs (Ruppert, 2015 ; Chen, 2024). La question de l'ouverture des données est aussi liée à celle de l'émergence des villes intelligentes, c'est-à-dire les villes dont le fonctionnement repose sur l'utilisation de technologies omniprésentes et productrices de données en temps réel, et dont l'économie et la gouvernance dépendent de l'innovation, de la créativité et de l'entrepreneuriat (Kitchin, 2014a ; Bernardin et Jeannot, *op. cit.*). Au-delà des enjeux de surveillance et de privatisation de l'espace public (Rogan, 2019), les villes intelligentes (*smart cities*) sont aussi vues comme le symptôme d'une nouvelle forme de gouvernamentalité, la gouvernamentalité « dirigée par les données » ou « gouvernamentalité algorithmique » (Henman, 2020).

- 26 Dans la lignée des travaux fondateurs sur la statistique d'État, ces recherches montrent comment la centralité des données et des calculs automatisés dans la conception des politiques publiques (Rieder et Simon, 2016) et dans les prises de décision transforme la manière dont sont perçus les citoyens et l'espace public, désormais réduits à des entités calculables. Plus récemment, les réponses des gouvernements face à la pandémie de COVID-19 ont donné lieu à de nombreuses études sur la gestion des populations et de leur santé grâce à des technologies de traçage et à l'analyse de données, d'une manière qui a souvent remis en question la liberté des citoyens et l'égalité dans l'accès aux soins (Lee, 2021 ; Schmidt, Roberts et Nwamaka, 2021 ; Pelizza *et al.*, 2021 ; Lupton, 2022). Finalement, dans un paysage technologique de plus en plus mondialisé et face à des enjeux géopolitiques et législatifs majeurs concernant les données et leur circulation, d'autres travaux explorent la question de la souveraineté des données, les États cherchant à avoir un meilleur contrôle sur les données produites par leurs citoyens et sur leur territoire (Irion, 2012). Enfin, les enjeux politiques sont également abordés à partir de la place de l'analyse et de l'exploitation des données par les partis politiques ou les États à des fins d'orientation, voire de manipulation, de l'opinion publique. Certains travaux s'intéressent à l'utilisation des données afin de rationaliser les efforts de ciblage de l'électorat lors des campagnes électorales (Theviot, 2016), dans un contexte où la place des données dans les actions de déstabilisation politique, à travers la diffusion de fausses informations et d'exploitation illégale des données personnelles, notamment à la suite du scandale de Cambridge Analytica, est largement discutée.
- 27 Sur le plan social, les conséquences de la mise en données sont abordées sous l'angle de la reproduction ou de l'aggravation des inégalités sociales et des discriminations. Un premier ensemble de recherches pointent le fait que les jeux de données mobilisés par les organisations, qu'ils soient antérieurs ou non à la numérisation, sont souvent biaisés et/ou incomplets, car hérités de logiques sociales inégalitaires préexistantes et d'une vision du monde patriarcale, raciste et colonialiste. Par exemple, certains travaux examinent les répercussions symboliques et matérielles de l'exploitation des données sur des groupes de population vulnérables et/ou minorisés. Un premier problème identifié est celui de l'absence de certaines catégories de population dans les données. Hargittai (2020) souligne que les groupes de personnes utilisant moins Internet et les médias sociaux, spécialement celles peu dotées en capitaux culturels et financiers, apparaissent peu dans les données massives comportementales récoltées en ligne destinées à la recherche ou au marketing, comme c'est le cas aussi des personnes âgées (Rosales et Fernández Ardèvol, 2019). Ces omissions font écho à celles relevées par les chercheurs et chercheuses qui s'intéressent aux données des recensements de la population et qui, faute de catégories adéquates, ont longtemps invisibilisé les citoyens

et citoyennes dont l'identité de genre ou l'orientation sexuelle ne correspondait pas aux normes dominantes (Bouk, 2022 ; Guyan, 2022).

- 28 D'autres catégories de population minorisées font bien l'objet d'une mise en données, mais d'une manière biaisée qui tend à reproduire des stéréotypes négatifs. C'est le cas des données étatiques concernant les peuples autochtones et les personnes noires, qui se concentrent sur les aspects négatifs de la vie de ces communautés, ce qui contribue à reproduire une vision déficitaire (*deficit narrative*) (Walter *et al.*, 2020 ; Gachoud, Naslund et Hansson, 2023 ; Watson-Daniels *et al.*, 2020). Par ailleurs, les efforts institutionnels pour développer une meilleure inclusion au moyen de nouvelles catégories, par l'exemple pour l'identification des personnes transgenres, peuvent reproduire des formes de violence (Hoffman, 2020), voire constituer un outil de surveillance qui pourrait les mettre en danger dans des contextes où les transidentités sont réprimées (D'Ignazio et Klein, *op. cit.*). En définitive, et au vu de l'importance des données et de la quantification dans la construction de la réalité, la circulation de ces jeux de données tronqués ou biaisés contribue à reproduire l'invisibilisation et les stéréotypes concernant les plus vulnérables, alimentant ainsi le terreau des discriminations.
- 29 Au-delà, d'autres travaux se concentrent sur les conséquences matérielles de cette inadéquation des données, en particulier lorsque leur traitement conduit au ciblage des groupes et des individus et à la prédiction de leurs comportements, à l'origine de prises de décision souvent discriminatoires dans les domaines du maintien de l'ordre et de la justice, de l'assistance sociale, de la santé ou encore des assurances. Plusieurs travaux fondateurs sur les conséquences des biais algorithmiques, à l'image de ceux de O'Neil (2016), de Noble (*op. cit.*) ou de Eubanks (2018), montrent que ces systèmes fonctionnent comme des boucles de rétroaction négatives, à l'instar de ceux utilisés par la police étasunienne (Benbouzid, 2018 ; Currie *et al.*, 2016), qui poussent les agents à cibler en priorité les quartiers défavorisés dont les populations sont historiquement plus criminalisées et surveillées, tandis que les quartiers plus aisés sont ignorés. Ces pratiques s'inscrivent dans une histoire plus ancienne de mobilisation des données par les institutions afin de pratiquer le *redlining*, une discrimination fondée sur le lieu de résidence (Koopman, 2020 ; Bakelmun et Shoenfeld, 2020).
- 30 Un second ensemble de recherches sur les conséquences de la *datafication* prolongent les réflexions sur la fracture des données (*data divide*), c'est-à-dire « la relation de pouvoir asymétrique entre ceux qui collectent, stockent et exploitent de grandes quantités de données et ceux qui sont ciblés par cette collecte » (Andrejevic, 2014, p. 1673, notre traduction). À l'image de l'étude d'Andrejevic qui a contribué à populariser cette expression, ces travaux se concentrent souvent sur les inégalités entre les internautes ordinaires et les organisations à partir de la question de la surveillance et de la vie privée. Il s'agit notamment de documenter le phénomène dit du paradoxe de la vie privée (*privacy paradox*), c'est-à-dire l'inadéquation entre les préoccupations exprimées par les internautes concernant l'exploitation de leurs données personnelles à des fins commerciales et le manque d'actions concrètes mises en place pour s'en protéger (Young et Quan-Haase, 2013). Dans les approches en *critical data studies*, ce « compromis » (Bourcier et De Filippi, 2016) des internautes entre la délégation de leurs données personnelles et la promesse d'une personnalisation des services est moins lu comme un phénomène relevant de la responsabilité individuelle que comme une forme d'impuissance dans un jeu de pouvoir à l'avantage des

entreprises. Ce phénomène doit par ailleurs être analysé au prisme des inégalités sociales et épistémiques entre les internautes (Kennedy, Steedman et Jones, 2020 ; Hoffmann, Lutz et Ranzini, 2024), et des émotions qu'il provoque (Grover, Idera et Ananny, 2025 ; Pink, Lanzeni et Horst, 2018).

- 31 De manière plus large, divers travaux portent sur le rapport des individus ordinaires aux données (Armstrong *et al.*, 2023 ; Burgess *et al.*, 2022), afin d'interroger leurs représentations et pratiques quotidiennes autour de l'auto-quantification de soi (*self-data*) (Lupton, 2020 ; Amelang, 2022), des recommandations algorithmiques (Siles, 2023 ; Bénistant, Chevret-Castellani et Labelle, 2024), de la géolocalisation (Dalton et Thatcher, 2019) et de l'interprétation des chiffres qui nous entourent, que ce soit sous forme de datavisualisation (Kennedy, Steedman et Jones, 2020) ou des données de suivi épidémiologiques lors de la pandémie de COVID-19 (Lee, 2023). Ces travaux ont pour intérêt de révéler la profondeur et la complexité des relations entre les individus ordinaires et les données (Feinberg, 2022) et d'apporter une lecture moins binaire du *data divide* en mettant l'accent sur les pratiques agentives de négociation, de réinterprétation, de résistance quotidienne, voire sur la constitution de « publics de données » actifs (Burgess *et al.*, 2025 ; Millerand *et al.*, 2025).
- 32 Les recherches comprennent également des travaux qui se concentrent, à une échelle plus macro, sur les inégalités liées à la mise en données entre les pays du Nord et ceux du Sud global. Un certain nombre de travaux se placent à l'interface entre les *critical data studies* et les études critiques sur le développement (Masiero, 2022). Ils étudient la manière dont les injonctions, les valeurs et les normes de la *datafication* telles que développées par les acteurs des pays du Nord tentent de s'imposer aux pays des Suds sans prise en compte de la pertinence et des bénéfices réels pour les populations, comme lorsque le *big data* et ses applications sont présentés comme un moyen de développement et d'*empowerment* (Arora, 2016) ; que des programmes de données ou science ouverte sont implantés sans tenir compte du contexte en termes de ressources (Bezuidenhout, Leonelli, Kelly et Rappert, 2017), de cultures de données (Lynch *et al.*, 2023) et de dynamiques de pouvoir locales (Gurstein, 2011). Ainsi, dans le contexte d'une relation qui demeure asymétrique entre ces pôles géographiques, ces politiques de mise en données contribuent à la fois à l'accélération de la *datafication* des Suds et à une potentielle réappropriation de ces données par des acteurs issus des pays du Nord au détriment des populations locales.
- 33 Un autre axe des travaux sur les conséquences de la mise en données s'intéresse aux mutations du capitalisme dans un contexte de mondialisation des industries numériques, en examinant les nouvelles chaînes de valeur basées sur le double mouvement de délocalisation et de sous-traitance, dans les pays du Sud global, des tâches sous-qualifiées du travail numérique (*digital labor*), au profit des grandes entreprises occidentales (Hung, 2024 ; Casilli, 2017). Ces microtâches qui consistent à « générer, annoter ou enrichir les données » (Le Ludec, Cornet et Casilli, 2023) assurées par des travailleurs de pays économiquement fragilisés d'Afrique et d'Asie permettent d'améliorer les systèmes algorithmiques utilisés dans les pays du Nord. L'un des buts des recherches critiques sur le sujet est alors de rendre visible ce travail, souvent ignoré des utilisateurs des services numériques (Anwar et Graham, 2020).
- 34 La mondialisation de l'économie des données est également étudiée à partir de son impact environnemental, dans un contexte où la consommation en ressources naturelles et en énergie exigée par le développement des systèmes automatisés

participe au réchauffement climatique et à la dégradation de la planète, dont les conséquences touchent majoritairement les pays vulnérables des Suds. Ces études se situent souvent au croisement entre *critical data studies* et *infrastructure studies*, ces dernières permettant d'aborder la dimension matérielle et (géo)politique des équipements liés à la circulation, au traitement et au stockage des données. Plusieurs recherches se penchent plus spécifiquement sur l'impact des centres de données (*data centers*) à une échelle mondiale (Brodie, 2024), mais aussi plus locale. C'est le cas, par exemple, des travaux de Hogan sur l'exploitation de l'eau par les serveurs de l'Agence nationale de sécurité (NSA) dans l'Utah aux États-Unis (Hogan, 2015) ou encore de ceux de Marquet (2018) sur les enjeux de l'implantation de centres de données en région parisienne.

Les conséquences sur les pratiques et les secteurs professionnels

- 35 À côté des travaux centrés sur la donnée et ses conséquences sociopolitiques, des études empiriques s'attellent à identifier et à analyser les conséquences de l'exploitation des données numériques dans un éventail très large de pratiques et de secteurs professionnels. Parmi ceux-ci, les plus nombreux se concentrent sur certains secteurs clés, comme la santé, l'éducation et la recherche scientifique⁷.
- 36 La santé est un secteur où les attentes suscitées par la *datafication* sont particulièrement fortes et où les données sont présentées comme un actif stratégique : sources d'avancées médicales pour les médecins, potentiel de connaissance pour les scientifiques, création de valeur économique pour les entreprises, opportunité de transparence pour les associations de patients, etc. (Peugeot, 2018). Sans surprise, de nombreux travaux en *critical data studies* se sont penchés sur l'étude critique des discours, imaginaires et promesses circulant sur les données de santé (Van Dijck et Peoll, 2016 ; Wamsley et Chin-Yee, 2021). Ces travaux montrent la prégnance d'un discours peu critique à l'endroit de la *datafication*, empreint des imaginaires associés au *big data*, où prolifèrent les métaphores sur la capture, l'exploitation et le partage de données au service du bien public, ce qui contribue à objectiver et à naturaliser la donnée de santé, une donnée pourtant caractérisée par l'hétérogénéité de ses contextes de production et d'usage (Stevens *et al.*, 2018 ; Cheung, 2020 ; Medina-Perea, Bates et Cox, 2024). Ce faisant, ces études montrent le pouvoir de ces métaphores et cadres de pensée sur l'organisation du secteur, en particulier depuis la pandémie, alors que l'essor des technologies numériques a catalysé de vastes changements dans la gouvernance mondiale de la santé, qui ont bénéficié surtout aux entreprises privées, par exemple à travers le développement d'applications et de plateformes en ligne (Wamsley et Chin-Yee, *op. cit.*).
- 37 D'autres études s'intéressent plus spécifiquement aux changements pour les patients et patientes, dans le cadre d'enquêtes empiriques menées sur des plateformes nationales de santé publique ou de services privés, avec des préoccupations centrées non seulement sur le gain d'autonomie potentiel pour les malades (grâce à l'accès aux données de santé à la maison, par exemple), mais aussi sur les risques pour la vie privée de la collecte de données à leur insu (Mahnke et Nielsen, 2021 ; Costa et Milne, 2022). Les travaux se développent sur les difficultés inhérentes à la conciliation d'attentes, d'imaginaires et de besoins différenciés autour des données de santé, avec des pistes vers le déploiement de stratégies de codesign des services et des plateformes, pour espérer concrétiser les promesses de retombées positives pour les patients et patientes,

qui restent encore largement à advenir (Donia et Shaw, 2021 ; Coutant et Vial, 2025 ; Harvey et Latzko-Toth, 2025).

- 38 L'éducation est un autre secteur où la mise en données promet de profondes transformations : le numérique y est déjà bien intégré et l'importance des systèmes de classement, notamment pour l'évaluation du rendement et de la performance, se prête particulièrement bien aux processus de *datafication*. Comme pour le secteur de la santé, la pandémie a précipité l'adoption de technologies et, avec elle, la production et la mise en circulation de données de toutes sortes. La controverse suscitée à l'été 2023 au sujet de la décision de la plateforme de téléconférence Zoom d'autoriser l'utilisation des données d'utilisation pour entraîner ses systèmes d'intelligence artificielle a remis sur le devant de la scène l'épineuse question de la protection des données personnelles, dans un secteur où ces outils sont désormais généralisés. Cette question alimente de nombreuses études en *critical data studies* sur les risques liés à la collecte systématique de données sur les élèves, de la maternelle aux études supérieures, sachant que la grande majorité des plateformes utilisées appartiennent aux grandes entreprises privées (Prinsloo, 2019 ; Prinsloo, Slade et Khalil, 2022).
- 39 À une échelle plus micro, des études empiriques analysent les plateformes numériques utilisées pour mesurer la performance scolaire, dans différents contextes nationaux (Hartong, 2021 ; Hartong et Förschler, 2019). Elles s'intéressent à la manière dont ces plateformes contribuent à définir ce qui constitue une « bonne » école, une école « performante » ou un élève « à risque », à travers les catégorisations et indicateurs essentiellement quantitatifs qu'elles proposent. Des études récentes se penchent notamment sur les usages qui sont faits des profils d'élèves « dataifiés » (Selwyn, Pangrazio et Cumb, 2022), qui construisent des subjectivités à partir d'uniques mesures de performance ou d'assiduité, mais qui se voient attribuer une valeur explicative globale du comportement des élèves. À ce titre, des pistes sont proposées pour recadrer le discours éducatif autour des données en les prenant comme un point de départ pour réfléchir aux améliorations possibles des pratiques professionnelles, plutôt que comme une source de connaissances guidant des décisions (*ibid.*).
- 40 La place grandissante dévolue aux données s'observe dans un autre secteur clé, la recherche scientifique, où l'étude des conséquences du traitement automatisé et à grande échelle de jeux de données sur la production des connaissances ne cesse d'alimenter les réflexions, et ce, depuis la vague du *big data* jusqu'à la diffusion récente des outils d'intelligence artificielle — des systèmes très gourmands en données. Les premiers travaux, stimulés par le contexte de « ruée vers l'or des données » qui a suivi la vague du *big data* (Felt, 2016), interrogent la manière dont l'accès à de vastes quantités de données numériques transforme les pratiques et postures épistémologiques des scientifiques (Plantin *et al.*, 2017 ; Kitchin, 2014b ; Leonelli, 2014 ; Markham, *op. cit.*). La fermeture de l'accès gratuit aux données par les plateformes en ligne et l'ajout de restrictions importantes pour les usages universitaires renouvellent actuellement cette question de l'accès aux données numériques à des fins de recherche scientifique (Birch, 2024).
- 41 Dans ces travaux qui interrogent les données dans le contexte scientifique, nous distinguons trois grandes problématiques récurrentes. La première est liée aux modalités et à l'éthique de « fabrication » des données et des connaissances à partir du moment où celles-ci sont produites en dehors de la sphère universitaire, en l'occurrence à partir de traces collectées par des plateformes commerciales gardant

secrets leurs algorithmes (Lagoze, 2014 ; Delias *et al.*, 2025). L'usage de méthodes numériques et computationnelles ainsi que l'intégration d'outils de visualisation de données dans le processus d'élaboration des connaissances alimentent également de nombreuses réflexions critiques (Cardon, 2012 ; Flyverbom et Murray, 2018 ; Flyverbom, Madsen et Rasche, 2017), avec des préoccupations liées aux niveaux de littératie de données (*data literacy*) chez les chercheuses et chercheurs formés en sciences humaines et sociales (Gray *et al.*, 2016). La deuxième problématique aborde, à une échelle plus macro, les liens d'interdépendance entre les milieux de la recherche et les entreprises productrices de données, notamment dans le domaine de la recherche biomédicale, mais aussi en sciences humaines et sociales (Richterich, 2018 ; Birch, 2024). La troisième s'intéresse à la réorganisation des disciplines et des expertises, sous l'action conjointe du développement de la science des données et de l'intégration de spécialistes de l'informatique et des données (*data scientists*) dans les équipes de recherche (Theviot et Treille, 2019 ; Ollion, 2016). Sur ce sujet, plusieurs travaux suggèrent d'incorporer des éléments issus des études critiques sur les données dans les programmes de formation des futurs scientifiques de données, qui restent centrés sur l'acquisition de compétences techniques en informatique et en statistique, ce qui limite leur compréhension des dimensions sociopolitiques et éthiques des données (Pène, 2020 ; Neff *et al.*, 2019 ; Stark et Hoffmann, 2019).

- 42 Enfin, nous distinguons deux problématiques transversales parcourant l'ensemble des recherches qui s'intéressent aux conséquences de la *datafication* sur les pratiques et secteurs professionnels. La première s'intéresse aux injonctions à développer une culture de la donnée au sein des organisations (Lehmans, 2017 ; Bates, 2017 ; Casemajor, 2025). Certaines études prennent la forme d'enquêtes empiriques sur la manière dont le personnel interagit avec les données et négocie l'instauration de nouvelles façons de faire, par exemple au sein d'administrations publiques engagées dans les données ouvertes (Millerand *et al.*, 2023 ; Millerand et Meunier, 2025). Dans un sens proche d'une littératie à développer tant à l'échelle individuelle qu'organisationnelle (Casemajor *et al.*, 2023), ces études se penchent sur les défis liés à l'intelligibilité des données et à l'intégration de processus de gestion des données dans les pratiques professionnelles existantes (Lehmans et Liquète, 2022 ; Lehmans et Capelle, 2023). La deuxième problématique transversale concerne les considérations éthiques, avec de nombreux questionnements touchant à la collecte de données sur les employés et aux risques de surveillance au travail, à l'évolution des codes d'éthique professionnelle, à la production de chartes éthiques, aux enjeux d'inclusion et de reproduction de biais dans les données, etc. (Metcalf et Crawford, 2016 ; Stark et Hoffmann, *op. cit.* ; Delias, 2025).

Les solutions de remplacement à la datafication

- 43 Face aux conséquences multiples de la *datafication*, qui transforment les logiques économiques, politiques et sociales à l'avantage des acteurs les plus puissants, un autre ensemble de recherches en *critical data studies* s'attellent à théoriser et à décrire les solutions de remplacement à ce phénomène. Certaines études se concentrent sur les efforts des organisations pour réguler les pratiques liées au traitement des données et corriger ou anticiper leurs potentiels effets négatifs, par exemple en analysant les discours des chartes éthiques dont se dotent les grandes entreprises du numérique, les agences et organismes nationaux et internationaux ainsi que les gouvernements et les collectivités territoriales (Jobin, Ience et Vayena, 2019 ; Verdi, 2023a) ou bien en

étudiant les différentes initiatives législatives contraignantes qui visent à encadrer l'usage des données (par exemple, le Règlement général sur la protection des données en Europe). Cependant, la majorité des recherches s'intéressent aux propositions issues de la société civile et se structurent autour de la notion de l'activisme de données (Gutiérrez, 2018). Ce terme désigne les initiatives citoyennes et mobilisations sociales visant à combattre les inégalités générées par l'exploitation des données, dans le but d'atteindre une forme de justice sociale autour des données (Taylor, 2017 ; Dencik *et al.*, 2022) caractérisée par une meilleure protection et une meilleure répartition du pouvoir entre les différents groupes sociaux, et ultimement de développer des savoirs et des imaginaires alternatifs de la mise en données (Baack, *op. cit.*). À la croisée des STS et de la sociologie des mouvements sociaux, l'activisme de données se situe à l'intersection des dimensions sociopolitiques propres aux actions citoyennes et des technologies, qui agissent comme médiatrices des enjeux démocratiques. Dans ce cadre, l'activisme de données est l'héritier de formes d'actions plus anciennes, comme le hacktivism ou le statactivism (Milan et Ven der Velden, 2016).

- 44 Beraldo et Milan (2019) proposent une typologie dont la dimension heuristique permet de cartographier les multiples initiatives relevant de l'activisme de données. Dans une première catégorie, les données et leur exploitation sont l'objet direct de la contestation, contre lesquelles la résistance s'organise (*data as stake*) : il s'agit par exemple de s'engager dans une association qui lutte contre la vidéosurveillance dans l'espace public. La seconde catégorie réunit les initiatives qui mobilisent les données et la science des données comme un moyen d'action pour promouvoir le changement social (*data as repertoires*), comme lorsque les citoyens utilisent les données ouvertes pour dénoncer l'existence de problèmes publics, par exemple dans les cas de pollution. Ces distinctions analytiques ne doivent pas cacher le fait que la plupart des initiatives d'activisme de données se situent en réalité sur un continuum entre ces pôles et empruntent souvent à plusieurs logiques en même temps.
- 45 Au sein des *critical data studies*, de nombreux travaux relèvent de la recherche-action ou de la recherche participative et proposent des outils et méthodes appropriables par les activistes et testés avec les citoyens. Parmi eux, un nombre important d'études s'intéressent à la manière dont les individus et collectifs peuvent « ouvrir la boîte noire » (Pasquale, 2015) que sont souvent les jeux de données et les systèmes techniques par lesquels ils sont traités par les gouvernements et les entreprises, afin de mettre au jour les biais et les rapports de pouvoir inégalitaires qui les sous-tendent. Il peut s'agir d'effectuer un « audit » technologique, dans la logique de la rétro-ingénierie : face à un système protégé par le secret industriel ou scientifique, certains experts mettent en place des protocoles de tests permettant de déduire leur fonctionnement. On peut citer les travaux sur les logiciels de reconnaissance faciale, dont les bases de données d'entraînement composées d'images de visages humains posent problème, soit parce que le manque de diversité des sujets entraîne une plus faible performance sur la reconnaissance des visages des femmes et des personnes non blanches (Buolamwini et Gebru, 2018), soit parce que ces données ont été récoltées d'une manière non éthique, comme dans le cas des portraits de youtubeurs trans étudiés par Keyes et Austin (2022).
- 46 De manière semblable à la démarche de Crawford dans son *Atlas de l'IA* (*op. cit.*) et dans un contexte où les technologies sont des composites transnationaux de plus en plus complexes, d'autres chercheurs et chercheuses proposent des méthodes pour

décortiquer la trajectoire des données afin de révéler les logiques de pouvoir qui président à leur création et à leur mise en circulation (Loukissas, 2017 ; Fordyce et Jethani, 2021 ; Flensburg et Lai, 2023). Ces méthodes sont souvent une première étape complémentaire d'actions où les données sont mobilisées comme outil de lutte. Il s'agit, le plus souvent, de créer des jeux de données citoyens alternatifs, ou « contre-données » (Ruppert, 2019), qui mettent en lumière les enjeux de protection de la vie privée ou qui représentent mieux la réalité vécue par les groupes de population sujets à des discriminations, à l'image du hackathon organisé par Currie *et al.* (*op. cit.*) pour compléter les jeux de données officiels, très parcellaires, répertoriant les homicides commis par la police de Los Angeles aux États-Unis. Ces contre-jeux de données, qui peuvent être mobilisés comme des leviers de connaissance et de plaidoyer, concernent par exemple les féminicides (D'Ignazio *et al.*, 2022), la représentation des personnes LGBT+ (Giesecking, 2018 ; Ruberg et Ruelos, 2020), l'environnement (Barreneche et Lombana-Bermudez, 2023), la justice raciale (Meng et DiSalvo, 2018) et les peuples autochtones (Rainie *et al.*, 2019). On peut également signaler des initiatives appartenant à d'autres traditions en sciences des données et de l'information, comme la contre-cartographie, qui consiste à utiliser « les outils de la cartographie institutionnelle et commerciale pour les appliquer de manière située et ascendante » (Dalton et Stallmann, 2018 ; Kidd, 2019), ou encore le « contre-archivage » (Ben-David, 2020 ; Currie et Paris, 2018), qui contourne la privatisation des données par les organisations.

- 47 À partir d'une approche différente, proche de la sociologie des mouvements sociaux, d'autres travaux se concentrent sur l'étude de groupes ou de mouvements de data-activistes, en analysant leurs valeurs et imaginaires, leurs répertoires d'action ou leurs dynamiques et structures sociales. On y retrouve, entre autres, des recherches sur des collectifs de scientifiques de données qui mettent bénévolement leur expertise au service des milieux communautaires et associatifs (Millette *et al.*, 2025), des groupes qui luttent contre la surveillance des citoyens (Kazansky et Milan, 2021 ; Anderson, 2022) ou encore des collectifs qui se sont mobilisés spontanément pour produire des outils de collecte et de suivi des données de santé lors de la pandémie de COVID-19 (Goëta, 2022 ; Kim *et al.*, 2024).
- 48 Enfin, un dernier ensemble de recherches s'attellent à théoriser, à mesurer et à opérationnaliser le concept de littératie des données. Celle-ci peut être définie comme la capacité d'accéder, d'interpréter, d'évaluer, de gérer, de manipuler et d'utiliser de manière critique et éthique les données (Calzada-Prado et Marzal, 2013) grâce à une connaissance des caractéristiques de la donnée, de son cycle de vie et des différents impacts engendrés par son usage, notamment en termes de sécurité et de protection de la vie privée (Crusoe, 2016) (Verdi, 2023b).
- 49 Déclinée en plusieurs sous-concepts en fonction du contexte d'application ou du type de données (*critical big data literacy*, Sander, 2020 ; *creative data literacy*, D'Ignazio, 2017 ; *data infrastructure literacy*, Gray, Gerlitz et Bounegru, 2018 ; *personal data literacy*, Pangrazio et Selwyn, 2019), la littératie des données propose une approche reposant sur l'autonomisation des individus et des communautés grâce à l'éducation et à la pédagogie critique (Pronzato et Markham, 2023), afin de tenter de combler l'asymétrie épistémique générée par la fracture des données et de créer les conditions nécessaires à la résistance face aux régimes de données inégalitaires.
- 50 En conclusion de cette partie sur les thèmes principaux qui structurent le champ des *critical data studies*, deux constats principaux ressortent. Tout d'abord, les premières études qui se développent à partir des années 2010 se concentrent sur l'étude des

discours autour de la *datafication* et des données, pour y déceler les visions et imaginaires les plus prégnants et la manière dont ceux-ci contribuent à orienter les décisions, à façonner les pratiques et à organiser la vie sociale. L'orientation critique des recherches s'attache à déconstruire les visions naturalisantes et essentialistes sur les données qui sont issues du *big data* (et donc de perspectives informatiques et gestionnaires) qui les présentent d'une manière apolitique, comme un actif stratégique au service du bien commun. Les études critiques sur les données révèlent alors les logiques de pouvoir à l'œuvre, leurs implications sur le plan économique, politique et social. De nombreux travaux s'orientent plus particulièrement vers l'analyse de la reproduction ou de l'aggravation des inégalités sociales et des discriminations à travers les processus de *datafication*, avec des préoccupations de justice sociale. Vont ensuite se développer des études empiriques qui contribueront à ancrer ces approches critiques dans une diversité de terrains d'études. Les études généralisantes sur les données laissent place à des enquêtes sur des dimensions particulières de la *datafication* (par exemple, l'étude de régimes de données dans un projet de ville intelligente, des cultures de données dans les organisations ou des formes d'activisme autour des données). Ces recherches prennent la forme d'études de cas ou encore s'attardent à des contextes spécifiques et concrets (par exemple, la pandémie de COVID-19 qui stimule la production et la circulation des données de santé) ou à des collectifs de travail ou d'activistes.

Conclusion

- 51 Cet article, qui vise à offrir une vision globale du champ des études critiques sur les données, a montré comment cet ensemble de travaux a émergé au croisement de différentes filiations théoriques et en réponse au positivisme naïf des études sur le *big data* qui dominaient les recherches au début des années 2010. Plaçant au centre de la démarche de recherche une approche sociotechnique et critique des données, ce champ d'études s'est structuré principalement autour de communautés scientifiques ancrées dans le monde anglophone. Les travaux francophones, même s'ils sont plus dispersés et s'ils ne revendiquent pas toujours l'étiquette de *critical data studies*, ont toutefois contribué au développement de travaux critiques sur les données, notamment avec des travaux fondateurs en sociologie de la quantification. L'article propose par ailleurs un bilan des thématiques qui structurent le champ, dont les recherches, qui s'ancrent au fil du temps de plus en plus dans l'analyse empirique des divers acteurs et pratiques autour des données, s'attellent à explorer les conséquences de la *datafication* sur l'ensemble des domaines de la vie sociale et des secteurs professionnels, mais aussi sur les résistances et les solutions de remplacement possibles aux régimes de données dominants.
- 52 Cette cartographie permet de faire émerger les apports principaux des *critical data studies* à l'étude des technologies en sciences humaines et sociales. Leur premier apport est d'établir que les données ne sont ni neutres ni brutes, au contraire de ce que certains discours peuvent laisser croire : elles sont inextricables des contextes de production et d'utilisation qui leur donnent vie ; elles sont ainsi façonnées par les enjeux politiques, culturels et économiques qui les traversent et témoignent de la réalité qu'elles documentent en charriant ces biais. Cependant, ces rapports de pouvoir demeurent asymétriques, dans un contexte où les organisations qui ont les moyens de

produire et d'exploiter à grande échelle les données qui influencent nos vies sociales jouissent de positions largement dominantes (Richterich, *op. cit.*). Le second apport est lié à l'objet même du champ : travailler sur les données revient à s'intéresser à des artefacts et à des dynamiques qui préexistent largement aux technologies numériques, comme le montrent les travaux sur la mise en nombre, mais aussi sur les pratiques de production et de partage des données scientifiques. Cette généalogie permet donc d'examiner les effets en propre du numérique : les données sur les populations, notamment, ont toujours été créées selon des normes et au service de stratégies politiques dominantes, à des fins de surveillance, de contrôle et de rationalisation (Mattelart, *op. cit.*). La numérisation de nos activités et les données massives qui en découlent viennent ainsi reproduire et intensifier des logiques préexistantes (Christin, *op. cit.*) ; ce constat permet donc de mettre à distance les discours déterministes qui parlent trop facilement de « révolution » technologique. Le troisième apport, lui aussi lié à l'objet des *critical data studies*, réside dans le fait que les données sont des artefacts essentiels au fonctionnement des technologies numériques qui dominent actuellement les investissements économiques, les discours et les pratiques, à savoir les technologies dites d'intelligence artificielle (IA), ce qui confirme l'importance stratégique des données et la pertinence de leur étude à partir d'un point de vue critique.

- 53 La montée des dispositifs d'IA générative et d'aide à la décision, comme ChatGPT ou les recommandations algorithmiques sur YouTube, constitue un nouveau terrain d'exploration pour les études critiques sur les données qui offrent une perspective particulièrement pertinente pour questionner ces outils inédits fonctionnant à partir d'une quantité faramineuse de données, mais aussi pour voir venir, et potentiellement prévenir, les écueils vers lesquels le recours à ces dispositifs pourrait nous mener. De même, la perspective critique et sociotechnique des *critical data studies* semble essentielle pour éclairer les décisions politiques en matière de gouvernance et de propriété des données, alors que la majorité des plateformes de médias sociaux appartiennent aux États-Unis et à la Chine — une question d'autant plus urgente que les affiliations politiques mondiales semblent se redéfinir. Enfin, les enjeux récents autour des risques de perte d'importants jeux de données scientifiques hébergés aux États-Unis, par exemple sur le climat, à la suite des élections de novembre 2024, rappellent la dimension éminemment politique des données et leur intérêt stratégique. Poursuivre le développement du champ des études critiques sur les données apparaît plus pertinent que jamais.

BIBLIOGRAPHIE

ANDREJEVIC Mark (2014), « The big data divide », *International Journal of Communication*, 8(17), p. 1673-1689.

ANWAR Mohammad Amir et Mark GRAHAM (2020), « Digital labour at economic margins: African workers and the global information economy », *Review of African Political Economy* 47(163), p. 95-105.

- AMELANG Katrin (2022), « (Not) safe to use: Insecurities in everyday data practices with period-tracking apps », dans Andreas HEPP, Julian JARKE et Leif KRAMP, *The ambivalences of data power: New perspectives in critical data studies*, Cham, Palgrave Macmillan, p. 297-321.
- ANDERSON Patrick (2022), *Cyberpunk ethics: Radical ethics for the digital age*, Londres, Routledge.
- ARMSTRONG Andrea, Jo BRIGGS, Wendy MONCUR, Daniel Paul CAREY, Emma NICOL et Burkhard SCHAFER (2023), « Everyday digital traces », *Big Data & Society*, 10(2), DOI : <https://doi.org/10.1177/20539517231213827>.
- ARORA Payal (2016), « Bottom of the data pyramid: Big data and the global south », *International Journal of Communication*, 10, p. 1681-1699.
- BAACK Stefan (2015), « Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism », *Big Data & Society*, 2(2), p. 1-11.
- BADOUARD Romain (2017), « *Open government, open data : l'empowerment citoyen en question* », dans Clément MABI, Jean-Christophe PLANTIN et Laurence MONNOYER-SMITH, *Ouvrir, partager, réutiliser. Regards critiques sur les données numériques*, Paris, Éditions de la Maison des sciences de l'homme, <https://books.openedition.org/editionsmsmh/9067>, page consultée le 4 décembre 2025..
- BAKELMUN Ashley et Sarah Jane SHOENFELD (2020), « Open data and racial segregation: Mapping the historic imprint of racial covenants and redlining on american cities », dans Scott HAWKEN, Honn HAN et Chris PETTIT (dir.), *Open cities – Open data: Collaborative cities in the information era*, Singapore, Springer, p. 57-83.
- BARRENECHE Carlos et Andres LOMBANA-BERMUDEZ (2023), « Another infrastructure is possible: Grassroots citizen sensing and environmental data justice in Colombia », *International Journal of Communication*, 17, p. 3638-3659.
- BATES Jo (2017), « Data cultures, power and the city », dans Rob KITCHIN, Tracey LAURIAULT et Gavin MCARDLE (dir.), *Data and the City*, Londres, Routledge, p. 189-200.
- BEER David (2017), « The social power of algorithms », *Information, Communication & Society*, 20(1), p. 1-13, DOI : <https://doi.org/10.1080/1369118X.2016.1216147>.
- BEN-DAVID Anat (2020), « Counter-archiving Facebook », *European Journal of Communication*, 35(3), p. 249-264.
- BENBOUZID Bilel (2018), « Quand prédire, c'est gérer. La police prédictive aux États-Unis », *Réseaux*, 211, p. 221-256.
- BÉNISTANT Alix, Christine CHEVRET-CASTELLANI et Sarah LABELLE (2024), « Quand les abonnés et abonnées travaillent leur profil. Netflix, algorithmes et promesses de découvrabilité », dans Colette BRIN et Véronique GUÈVREMONT (dir.), *IA, culture et médias*, Presses de l'Université de Laval, <https://hal.science/hal-04322289v1>, page consultée le 4 décembre 2025.
- BERALDO Davide et Stefania MILAN (2019), « From data politics to the contentious politics of data », *Information, Communication & Society*, 22(7), p. 955-971.
- BERNARDIN Stève et Gilles JEANNOT (2019), « La ville intelligente sans les villes ? », *Réseaux*, 6(218), p. 93-107, <https://www.cairn.info/revue-reseaux-2019-6-page-93.htm>.
- BEZUIDENHOUT Louise, Sabina LEONELLI, Ann H. KELLY et Brian RAPPERT (2017), « Beyond the digital divide: Towards a situated approach to open datas », *Science and Public Policy*, 44(4), p. 464-475.

- BIGO Didier, Engin ISIN et Evelyn RUPPERT (2019), *Data politics. Worlds, subjects, rights*, Londres, Routledge.
- BIRCH Kean (2023), *Data Enclaves*, Cham, Springer Nature.
- BIRCH Kean (2024), « Assetization as a mode of techno-economic governance: Knowledge, education and personal data in the un's system of national accounts », *Economy & Society*, 53(1), p. 15-38.
- BOUK Dan (2022), *Democracy's Data: The Hidden Stories in The U.S. Census and How to Read Them*, New York, MCD/Farrar, Straus and Giroux.
- BOURCIER Danièle et Primavera DE FILIPPI (2016), *Open data & big data : nouveaux défis pour la vie privée*, Paris, Mare & Martin.
- BOWKER Geoffrey C., Karen BAKER, Florence MILLERAND et David RIBES (2010), « Toward information infrastructure studies », dans Jeremy HUNSINGER, Lisbeth KLAstrup et Matthew ALLEN (dir.), *International Handbook of Internet Research*, Dordrecht, Springer, p. 97-117.
- BOYD danah et Kate CRAWFORD (2012), « Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon », *Information, Communication & Society*, 15(5), p. 662-679.
- BRODIE Patrick (2024), « Smarter, greener extractivism: Digital infrastructures and the harnessing of new resources », *Information, Communication & Society*, 28(6), p. 1061-1080.
- BROUSSARD Meredith (2023), *More than a glitch: Confronting race, gender, and ability bias in tech*, Cambridge, MIT Press.
- BULLICH Vincent et Viviane CLAVIER (2018), « Production des données, "production de la société" », *Les Enjeux de l'information et de la communication*, 2(19), p. 51-64.
- BUOLAMWINI Joy et Timnit GEBRU (2018), « Gender shades: Intersectional accuracy disparities in commercial gender classification », *Proceedings of Machine Learning Research*, 81, p. 1-15.
- BURGESS Jean (2021), « Platform Studies », dans Stuart CUNNINGHAM et David CRAIG (dir.), *Creator Culture: An Introduction to Global Social Media Entertainment*, New York, New York University Press, p. 21-38.
- BURGESS Jean, Kate ALBURY, Anthony McCOSKER et Rowan WILKEN (2022), *Everyday data cultures*, Cambridge, Polity.
- BURGESS Jean, Kath ALBURY, Anthony MCCOSTER et Rowan WILKEN (2025), « Les publics de données quotidiens », dans Florence MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l'Université de Montréal, p. 155-182.
- CALZADA-PRADO Javier et Miguel Ángel MARZAL (2013), « Incorporating data literacy into information literacy programs: Dore competencies and contents », *Libri*, 63(2), DOI : <https://doi.org/10.1515/libri-2013-0010>.
- CARDON Dominique (2012), « Regarder les données », *Multitudes*, 49, p. 138-142.
- CARDULLO Paolo et Rob KITCHIN (2019), « Being a "citizen" in the smart city: Up and down the scaffold of smart citizen participation in Dublin, Ireland », *GeoJournal*, 84, p. 1-13, DOI : <https://doi.org/10.1007/s10708-018-9845-8>.
- CASEMAJOR Nathalie (2025), « Les cultures de données. Genèse d'une notion, enjeux épistémologiques et implications pratiques », dans Florence MILLERAND, Alexandre COUTANT,

Guillaume LATZKO-TOTH, Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l'Université de Montréal, p. 93-110.

CASEMAJOR Nathalie, Guillaume SIROUS, Guy BELLAVANCE, Nathalie VACHON, François R. DERBAS THIBODEAU, Alexandra BEAUSÉJOUR-VILLEMAIRE et Alexis MORIN (2023), *Gouvernance des données d'usage dans les bibliothèques et archives : Sondage sur les perceptions des usagers de Bibliothèque et Archives nationales du Québec*, Montréal, INRS – Urbanisation Culture Société.

CASILLI Antonio A. (2017), « Global digital culture: Digital labor studies go global. Toward a digital decolonial turn », *International Journal of Communication*, 11, p. 3934-3954.

CHEN Tzu-Ting-Yi (2024), « Strong or thin digital democracy? The democratic implications of Taiwan's open government data policy in the 2010s », *Big Data & Society*, 11(4), DOI : <https://doi.org/10.1177/20539517241296>.

CHEUNG Sarah (2020), « Disambiguating the Benefits and Risks from Public Health Data in the Digital Economy », *Big Data & Society*, 7(1), DOI : <https://doi.org/10.1177/2053951720933924>.

CHEVALLIER Jacques (2018), « Vers l'état-plateforme ? », *Revue française d'administration publique*, 167(3), p. 627-637.

CHRISTIN Angèle (2020), « What data can do: A typology of mechanisms », *International Journal of Communication*, 14, p. 1115-1134.

COSTA Alessia et Richard MILNE (2022), « Understanding “passivity” in digital health through imaginaries and experiences of coronavirus disease 2019 contact tracing apps », *Big Data & Society*, 9(1), DOI : <https://doi.org/10.1177/20539517221091138>.

COULDRY Nick et Ulises A. MEJIAS (2018), « Data colonialism: Rethinking big data's relation to the contemporary subject », *Television & New Media*, 20(4), p. 336-349.

COUTANT Alexandre et Stéphane VIAL (2025), « Le codesign d'un service numérique de santé mentale. Concilier les imaginaires et les attentes autour des données », dans Florence MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l'Université de Montréal, p. 75-92.

CRAIN Matthew (2018), « The limits of transparency: Data brokers and commodification », *New Media & Society*, 20(1), DOI : <https://doi.org/10.1177/1461444816657096>.

CRAWFORD Kate (2021), *The atlas of AI*, New Haven, Yale University Press, DOI : <https://doi.org/10.2307/j.ctv1nh3kvb>.

CRAWFORD Kate, Mary L. GRAY et Kate MILTNER (2014), « Critiquing big data: Politics, ethics, epistemology », *International Journal of Communication*, 8, p. 1663-1672.

CRÉMIER Lucile, Maude BONENFANT et Laura Iseult LAFRANCE (2019), « Raw data or hypersymbols? Meaning-making with digital data, between discursive processes and machinic procedures », *Semiotica*, 230, p. 189-212.

CRUSOE David (2016), « Data literacy defined pro populo: To read this article, please provide a little information », *The Journal of Community Informatics*, 12(3), DOI : <https://doi.org/10.15353/joci.v12i3.3276>.

CURRIE Morgan E., Britt S. PARIS, Irene PASQUETTO et Jennifer PIERRE (2016), « The conundrum of police officer-involved homicides: Counter-data in Los Angeles county », *Big Data & Society*, 3(2), DOI : <https://doi.org/10.1177/2053951716663566>.

CURRIE Morgan E. et Britt S. PARIS (2018), « Back-ups for the future: Archival practices for data activism », *Archives and Manuscripts*, 46(2), p. 124-142.

DAGIRAL Éric, Christian LICOPPE, Olivier MARTIN et Anne-Sophie PHARABOD (2019), « Le quantified self en question(s). Un état des lieux des travaux de sciences sociales consacrés à l'automesure des individus », *Réseaux*, 4(216), p. 17-54.

DALTON Craig et Jim THATCHER (2014), « What does a critical data studies look like, and why do we care? Seven points for a critical approach to “Big Data” », *Society & Space*, <https://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care>, page consultée le 30 novembre 2025.

DALTON Craig M. et Tim STALLMAN (2018), « Counter-mapping data science », *Canadian Geographer*, (62)1, p. 93-101.

DALTON Craig et Jim THATCHER (2019), « Seeing by the Starbucks: The social context of mobile maps and users' geographic knowledges », *Cartographic Perspectives*, 92, DOI : <https://doi.org/10.14714/CP92.1447>.

DALTON Craig, Lindsay TAYLOR et Jim THATCHER (2016), « Critical data studies: A dialog on data and space », *Big Data & Society*, 3(1), DOI : <https://doi.org/10.1177/2053951716648346>.

DELIAS Lucie (2025), « L'inclusion des groupes minorisés dans les politiques publiques », dans Florence MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l'Université de Montréal, p. 141-154.

DELIAS Lucie, Claudine BONNEAU, Guillaume LATZKO-TOTH et Florence MILLERAND (2025), « Traces numériques, données et preuves en sciences humaines et sociales », dans Julien PRUD'HOMME et Molly KAO (dir.), *Preuve et sciences*, Montréal, Presses de l'Université de Montréal.

DENCIK Lina, Emiliano TRERÉ, Arne HINTZ et Joanna REEDER (2022), *Data justice*, Londres, Sage.

DENIS Jérôme (2018), *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines.

DESROSIÈRES Alain (2000), *La politique des grands nombres : histoire de la raison statistique*, Paris, La Découverte.

D'IGNAZIO Catherine (2017), « Creative data literacy: Bridging the gap between the data-haves and data-have nots », *Information Design Journal*, 23(1), p. 6-18.

D'IGNAZIO Catherine et Lauren F. KLEIN (2020), *Data feminism*, Cambridge, MIT Press.

D'IGNAZIO Catherine, Isadora CRUXÊN, Helena SUÁREZ VAL, Angeles MARTINEZ CUBA, Mariel GARCÍA-MONTES, Silvana FUMEGA, Harini SURESH et Wonyoung SO (2022), « Femicide and counterdata production: Activist efforts to monitor and challenge gender-related violence », *Patterns*, 3(7), DOI : <https://doi.org/10.1016/j.patter.2022.100530>.

DONIA Joseph et James A. SHAW (2021), « Co-design and ethical artificial intelligence for health: An agenda for critical research and practice », *Big Data & Society*, 8(2), DOI : <https://doi.org/10.1177/205395172110652>.

DURAND Cédric (2021), *Techno-féodalisme : critique de l'économie numérique*, Paris, Édition Zones.

EUBANKS Virginia (2018), *Automating inequality: How high-tech tools profile, police, and punish the poor*, New York, St. Martin's Press.

- FEINBERG Melanie (2022), *Everyday adventures with unruly data*, Cambridge, MIT Press.
- FELT Mylynn (2016), « Social media and the social sciences: How researchers employ big data analytics », *Big Data & Society*, 3(1), DOI : <https://doi.org/10.1177/2053951716645828>.
- FLENSBURG Sofie et Signe Sophus LAI (2023), « Follow the data! A strategy for tracing infrastructural power », *Media and Communication*, 11(2), p. 319-329.
- FLICHY Patrice (2021), « Le gouvernement des données : une nouvelle forme de gouvernance ? », *Réseaux*, 3(227), p. 107-130, <https://www.cairn.info/revue-reseaux-2021-3-page-107.htm>, page consultée le 30 novembre 2025.
- FLICHY Patrice et Olivier PARASIE (2013), « Le savant, l'expert et le politique », *Revue française de science politique*, 63(1), p. 7-31, DOI : <https://doi.org/10.3917/rfsp.631.0007>.
- FLYVERBOM Mikkel et Julian MURRAY (2018), « Datastructuring — Organizing and curating digital traces into action », *Big Data & Society*, 5(2), DOI : <https://doi.org/10.1177/2053951718799114>.
- FLYVERBOM Mikkel, Anders Koed MADSEN et Andreas RASCHE (2017), « Big data as governmentality in international development: Digital traces, algorithms, and altered visibilities », *The Information Society*, p. 35-42.
- FORDYCE Robbie et Suneel JETHANI (2021), « Critical data provenance as a methodology for studying how language conceals data ethics », *Continuum*, 135(5), p. 775-787.
- FUCHS Christian (2019), « What is critical digital social research? Five reflections on the study of digital society », *Journal of Digital Social Research*, 1(1), DOI : <https://doi.org/10.33621/jdsr.v1i1.7>.
- GACHOUD Vendela, Anna NASLUND et Karin HANSSON (2023), « Sami traces: Diversity and curatorial workarounds in image archives », dans Danilo GIGLITTO, Luigina CIOLFI, Eleanor LOCKLEY et Eirini KALDELI (dir.), *Digital approaches to inclusion and participation in cultural heritage*, Londres, Routledge, p. 181-206.
- GERRARD Ysabel et Jo BATES (2019), « Data power », *Online Information Review*, 43(6), DOI : <https://doi.org/10.1108/oir-10-2019-396>.
- GIESEKING Jen Jack (2018), « Size matters to lesbians, too: Queer Feminist interventions into the scale of big data », *The Professional Geographer*, 70(1), p. 150-156.
- GILLESPIE Tarleton (2014), « The relevance of algorithms », *Media Technologies*, 3, p. 167-194.
- GILLESPIE Tarleton (2015), « Platforms intervene », *Social Media + Society*, 1(1), DOI : <https://doi.org/10.1177/2056305115580479>.
- GILLESPIE Tarleton (2018), *Custodians of the Internet*, New Haven, Yale University Press.
- GITELMAN Lisa (dir.) (2013), « Raw data » is an oxymoron, Cambridge, MIT Press.
- GOËTA Samuel (2022), « La pandémie de Covid-19 : un moment charnière pour l'ouverture des données en France », *Statistique et société*, 10(2), p. 9-29.
- GOËTA Samuel et Clément MABI (2014), « L'open data peut-il (encore) servir les citoyens ? », *Mouvements*, 79(3), p. 88-91.
- GRAY Jonathan, Carolin GERLITZ et Liliana BOUNEGRU (2018), « Data infrastructure literacy », *Big Data & Society*, 5(2), DOI : <https://doi.org/10.1177/2053951718786316>.

- GRAY Jonathan, Liliana BOUNEGRU, Stefania MILAN et Paolo CIUCCARELLI (2016), « Ways of seeing data: Toward a critical literacy for data visualizations as research objects and research devices », dans Sebastian KUBITSCHKO et Anne KAUN (dir.), *Innovative methods in media and communication research*, Cham, Springer International Publishing, p. 227-251.
- GROVER Rohan, Josh IDERA et Mike ANANNY (2025), « Data disaffection: Toward a relational and affective understanding of datafication », *New Media & Society*, DOI : <https://doi.org/10.1177/14614448241310901>.
- GURSTEIN Michael B. (2011), « Open data: Empowering the empowered or effective data use for everyone? », *First Monday*, 16(2), DOI : <https://doi.org/10.5210/fm.v16i2.3316>.
- GUTIÉRREZ Miren (2018), *Data activism and social change*, New York, Palgrave Macmillan.
- GUYAN Kevin (2022), *Queer data: Using gender, sex and sexuality data for action*, Londres, Bloomsbury Academic.
- HARGITTAI Eszter (2020), « Potential bias in big data: Omitted voices on social media », *Social Media + Society*, 6(3), DOI : <https://doi.org/10.1177/0894439318788322>.
- HARTONG Sigrid (2021), « The power of relation-making: insights into the production and operation of digital school performance platforms in the US », *Critical Studies in Education*, 62(1), p. 34-49.
- HARTONG Sigrid et Annina FÖRSCHLER (2019), « Opening the black box of data-based school monitoring: Data infrastructures, flows and practices in state education agencies », *Big Data & Society*, 6(1), DOI : <https://doi.org/10.1177/2053951719853311>.
- HARVEY Maxime et Guillaume LATZKO-TOTH (2025), « La représentation des citoyens dans une plateforme de données en santé », dans Florence MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l'Université de Montréal, p. 59-74.
- HENMAN Paul (2020), « Improving public services using artificial intelligence: Possibilities, pitfalls, governance », *Asia Pacific Journal of Public Administration*, 42(4), p. 209-221.
- HENNEGUELLE Aurélien et Jason JATTEAU (2021), « Les économistes face aux données », *Genèses*, 122(1), p. 10-33, DOI : <https://doi.org/10.3917/gen.122.0010>.
- HEPP Andreas, Julia JARKE et Leif KRAMP (dir.) (2022), *New perspectives in critical data studies: The ambivalences of data power*, Cham, Springer International Publishing.
- HOFFMANN Anna Lauren (2020), « Terms of inclusion: Data, discourse, violence », *New Media & Society*, 23(12), DOI : <https://doi.org/10.1177/1461444820958725>.
- HOFFMANN Chistina Pieter, Christoph LUTZ et Giulia RANZINI (2024), « Inequalities in privacy cynicism: An intersectional analysis of agency constraints », *Big Data & Society*, 11(1), DOI : <https://doi.org/10.1177/20539517241232629>.
- HOGAN Mél (2015), « Data flows and water woes: The Utah data center », *Big Data & Society*, 2(2), DOI : <https://doi.org/10.1177/2053951715592429>.
- HUNG Kai-Hsin (2024), « Artificial intelligence as planetary assemblages of coloniality: The new power architecture driving a tiered global data economy », *Big Data & Society*, 11(4), DOI : <https://doi.org/10.1177/20539517241289443>.
- ILIADIS Andrew et Federica RUSSO (2016), « Critical data studies: An introduction », *Big Data & Society*, 3(2), DOI : <https://doi.org/10.1177/2053951716674238>.

- IRION Kristina (2012), « Government cloud computing and national data sovereignty », *Policy & Internet*, 4(3-4), p. 40-71.
- JOBIN Anna, Marcello IENCA et Effy VAYENA (2019), « The global landscape of ai ethics guidelines », *Nature Machine Intelligence*, 1(9), p. 389-399.
- KAZANSKY Becky et Stefania MILAN (2021), « “Bodies not templates”: Contesting dominant algorithmic imaginaries », *New Media & Society*, 23(2), DOI : <https://doi.org/10.1177/1461444820929316>.
- KENNEDY Helen et Jo BATES (2017), « Data power in material contexts », *Television and New Media*, 18(8), p. 701-705.
- KENNEDY Helen, Robin STEEDMAN et Rhianna JONES (2020), « Approaching public perceptions of datafication through the lens of inequality: A case study in public service media », *Information, Communication & Society*, 24(12), p. 1776-1794.
- KEYES Os et Jevan AUSTIN (2022), « Disability and data justice: Understanding accessibility as a data issue », *Big Data & Society*, 9(1), DOI : <https://doi.org/10.3233/EFI-230022>.
- KIDD Dorothy (2019), « Extra-activism: Counter-mapping and data justice », *Information, Communication & Society*, 22(7), p. 854-970.
- KIM Youngrim, Megan FINN, Amelia ACKER, Bidisha CHADHURI, Stacey WEDLAKE, Ryan ELLIS et Janaki SRINIVASAN (2014), « Epistemologies of missing data: COVID dashboard builders and the production and maintenance of marginalized COVID data », *Big Data & Society*, 11(2), DOI : <https://doi.org/10.1177/20539517241259666>.
- KITCHIN Rob (2014a), « The real-time city? Big data and smart urbanism », *GeoJournal*, 79(1), DOI : <https://doi.org/10.1007/s10708-013-9516-8>.
- KITCHIN Rob (2014b), « Big data, new epistemologies and paradigm shifts », *Big Data & Society*, 1(1), DOI : <https://doi.org/10.1177/2053951714528481>.
- KITCHIN Rob (2022), *The data revolution: A critical analysis of big data, open data & data infrastructures*, Londres, Sage.
- KITCHIN Rob (2025), *Critical data studies: An A to Z guide to concepts and methods*, Cambridge, Polity Press.
- KITCHIN Rob et Tracey LAURIAULT (2014), « Towards critical data studies: Charting and unpacking data assemblages and their work », *The Programmable City Working Paper 2*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112, page consultée le 30 novembre 2025.
- KITCHIN Rob et Gavin McARDLE (2016), « What makes big data, big data? Exploring the ontological characteristics of 26 datasets », *Big Data & Society*, 3(1), p. 1-10.
- KNOX Hannah et Evelina GAMBINO (2023), « Infrastructure », dans *The open encyclopedia of anthropology*, DOI : <http://doi.org/10.29164/23infrastructure>.
- KOOPMAN Colin (2020), *How we became our data: A genealogy of the informational person*, Chicago, University of Chicago Press.
- LAGOZE Clifford (2014), « Big data, data integrity, and the fracturing of the control zone », *Big Data & Society*, 1(2), DOI : <https://doi.org/10.1177/2053951714558281>.
- LATOUR Bruno (2007), « Le “pédofil” de Boa Vista – Montage photo-philosophique », dans *Petites leçons de sociologie des sciences*, Paris, La Découverte, p. 171-225.

LATZKO-TOTH Guillaume, Claudine BONNEAU et Mélanie MILLETTE (2017), « Small data, thick data: Thickening strategies for trace-based social media research », dans Anabel QUAND-HASSE et Luke SLOAN (dir.), *The SAGE handbook of social media research methods*, Londres, Sage, p. 199-214.

LATZKO-TOTH Guillaume, Claudine BONNEAU et Mélanie MILLETTE (2020), « La densification des données : revaloriser la recherche qualitative à l'ère des données massives », dans Mélanie MILLETTE, Florence MILLERAND, David MYLES et Guillaume LATZKO-TOTH (dir.), *Méthodes de recherche en contexte numérique : une orientation qualitative*, Montréal, Presses de l'Université de Montréal, p. 181-194.

LEE Jeehyun Jenny (2023), « Vital dataveillance: Investigating data in exchange for vitality through South Korea's COVID-19 technogovernance », *Communication, Culture and Critique*, 15(4), p. 499-506.

LEE Benjamin (2021), « Compounded mediation: A data archaeology of the newspaper navigator dataset », *Digital Humanities Quarterly*, 15(14), DOI : <https://doi.org/10.17613/k9gt-6685>.

LEHMANS Anne (2017), « Données ouvertes et redéfinition de la culture de l'information dans les organisations », *Communication et organisation. Revue scientifique francophone en communication organisationnelle*, 51, p. 15-26.

LEHMANS Anne et Camille CAPELLE (2023), « Le cadre de l'expérience des données en éducation : gouvernance, représentations et intelligibilité des données dans l'éducation nationale », *Communication et organisation. Revue scientifique francophone en communication organisationnelle*, 64, p. 33-49.

LEHMANS Anne et Vincent LIQUÈTE (2022), « Des littératies aux approches culturelles du numérique : l'exemple de la culture des données dans le champ de l'éducation aux médias et à l'information (EMI) », *Approches théoriques en information-communication*, 5(2), p. 35-46.

LE LUDEC Clément, Maxim CORNET et Antoinio A. CASILLI (2023), « The problem with annotation. Human labour and outsourcing between France and Madagascar », *Big Data & Society*, 10(2), DOI : <https://doi.org/10.1177/20539517231188723>.

LEONELLI Sabina (2014), « What difference does quantity make? On the epistemology of big data in biology », *Big Data & Society*, 1(1), DOI : <https://doi.org/10.1177/2053951714534395>.

LINGREN Simon (2023), *Handbook of critical studies of artificial intelligence*, Cheltenham, Edward Elgar.

LOUKISSAS Yanni Alexander (2017), « Taking big data apart: Local reading of composite media collections », *Information, Communication & Society*, 20(5), p. 651-664.

LUKA Mary Elizabeth et Mélanie MILLETTE (2018), « (Re)framing big data: Activating situated knowledges and a feminist ethics of care in social media research », *Social Media + Society*, 4(2), DOI : <https://doi.org/10.1177/2056305118768297>.

LUPTON Deborah (2015), *Digital sociology*, Londres, Routledge, DOI : <https://doi.org/10.4324/9781315776880>.

LUPTON Deborah (2020), « Data mattering and self-tracking: What can personal data do? », *Continuum*, 34(1), p. 1-13.

LUPTON Deborah (2022), « How do data come to matter? Living and becoming with personal data », *Social Media + Society*, 8(1), DOI : <https://doi.org/10.1177/20563051221086967>.

LYCETT Mark (2013), « "Datafication": Making sense of (big) data in a complex world », *European Journal of Information Systems*, 22(4), p. 381-386.

LYNCH Renee, Jason C. YOUNG, Chris JOWAISAs, Joel SAM, Stanley BOAKYE-ACHAMPONG, Maria GARRIDO et Chris ROTHSCCHILD (2023), « “The tears don’t give you funding”: Data neocolonialism in development in the global south », *Third World Quarterly*, 44(5), DOI : <https://doi.org/10.1080/01436597.2023.2166482>.

MAHNKE Martina Skrubbeltrang et Mikka NIELSEN (2021), « E-health platforms as user-data communication: Examining patients’ struggles with digital health data », *Nordicom Review*, 42(4), p. 45-58.

MARCETIC Hana et Jan NOLIN (2022), « Feminist data studies and the emergence of a new feminist knowledge domain », *First Monday*, 27(7), DOI : <https://doi.org/10.5210/fm.v27i7.12295>.

MARKHAM Annette N. (2013), « Undermining “data”: A critical examination of a core term in scientific inquiry », *First Monday*, 18(10), DOI : <https://doi.org/10.5210/fm.v18i10.4868>.

MARKHAM Annette, Katrin TIIDENBERG et Andrew HERMAN (2018), « Doing ethics in the era of big data research », numéro spécial de *Social Media + Society*, 4(2), DOI : <https://doi.org/10.1177/2053951718766651>.

MARQUET Clément (2018), « Ce nuage que je ne saurais voir. Promouvoir, contester et réguler les data centers à Plaine Commune », *Tracés. Revue de science humaines*, 2(35), p. 75-98.

MARTIN Olivier (2020), *L’empire des chiffres*, Paris, Armand Colin.

MASIERO Silvia (2022), « ICT4D and critical data studies: A joint research agenda », *ACM International Conference Proceeding Series*, Association for Computing Machinery, DOI : <https://doi.org/10.1145/3572334.3572388>.

MATTELART Armand (2019), « Pour une archéologie du culte du nombre », dans Éric GEORGE (dir.), *Numérisation de la société et enjeux sociopolitiques 1*, Montréal, ISTE, p. 19-30.

MAYER-SCHÖNBERGER Viktor et Kenneth CUKIER (2013), « Big data: A revolution that will transform how we live, work, and think », *American Journal of Epidemiology*, 179(9), p. 1143-1144.

MEDINA-PEREA Itzelle, Jo BATES et Andrew COX (2024), « ‘Data saves lives’: Ideational-material drivers of health data journeys in the UK », *Big Data & Society*, 11(4), DOI : <https://doi.org/10.1177/20539517241296056>.

MEJIAS Ulises A. et Nick COULDRY (2024), *Data Grab. The New Volonialism of Big Tech and How to Fight Back*, Londres, WH Allen.

MEJIAS Ulises Ali, Sarah McNEALY et Monica MICELI (à paraître, 2026), *De Gruyter handbook of critical data studies*, Berlin, De Gruyter.

MÉNARD Mathieu, Alexandre MONDOUX, Maxime OUELLET et Mathieu BONENFANT (2018), « Big data, gouvernementalité et industrialisation des médiations symboliques et politico-institutionnelles », *Interfaces numériques*, 5(2), p. 31-41.

MENG Amanda et Carl DiSALVO (2018), « Grassroots resource mobilization through counter-data action », *Big Data & Society*, 5(2), DOI : <https://doi.org/10.1177/2053951718796862>.

METCALF Jacob et Kate CRAWFORD (2016), « Where are human subjects in big data research? The emerging ethics divide », *Big Data & Society*, 3(1), DOI : <https://doi.org/10.1177/2053951716650211>.

MILAN Stefania et Lonneke VAN DER VELDEN (2016), « The alternative epistemologies of data activism », *Digital Culture & Society*, 2(2), p. 57-74.

- MILLERAND Florence, Lucie DELIAS, Alexandre COUTANT et Marie-Soleil FORTIER (2023), « À la recherche du citoyen “ordinaire”, les publics imaginés de l’ouverture des données publiques au niveau municipal », *Les Enjeux de l’information et de la communication*, 234(1), p. 109-129.
- MILLERAND Florence et Sarah MEUNIER (2025), « Les publics “impensés” des données ouvertes. Quand les fournisseurs désignés sont aussi les usagers », dans Florence MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l’Université de Montréal, p. 25-40.
- MILLERAND Florence, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (2025), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l’Université de Montréal.
- MILLETTE Mélanie (2023), « Épistémologie et méthodologie féministes pour la recherche en contexte numérique », *Communication*, 2(40), DOI : <https://doi.org/10.4000/communication.17969>.
- MILLETTE Mélanie, Florence MILLERAND et Nadia SERAIOCCO (2025), « Data for Good. La science des données au service des publics minorisés et des milieux communautaires », dans Florence MILLERAND, Alexandre COUTANT, Guillaume LATZKO-TOTH et Mélanie MILLETTE (dir.), *Les publics de données. Penser la datafication de la société*, Montréal, Presses de l’Université de Montréal, p. 125-140.
- NEFF Gina, Anissa TANWEER, Brittany FIORE-GARTLAND et Laura OSBURN (2017), « Critique and contribute: A practice-based framework for improving critical data studies and data science », *Big Data*, 5(2), p. 85-97.
- NOBLE Safiya Umoja (2018), *Algorithms of oppression: How search engines reinforce racism*, New York, NYU Press.
- OLLION Étienne (2016), « L’abondance et ses revers : Big data, open data et recherches sur les questions sociales », *Informations sociales*, 191, p. 70-79.
- O’NEIL Cathy (2016), *Weapons of math destruction: How big data increases inequality and threatens democracy*, Londres, Penguin.
- OUELLET Maxime, Mars MÉNARD, Maude BONENFANT et André MONDOUX (2015), « Big Data et quantification de soi : la gouvernementalité algorithmique dans le monde numériquement administré », *Canadian Journal of Communication*, 40(4), DOI : <https://doi.org/10.22230/cjc.2015v40n4a2869>.
- PANGRAZIO Luci et Neil SELWYN (2019), « “Personal data literacies”: A critical literacies approach to enhancing understandings of personal digital data », *New Media & Society*, 21(2), DOI : <https://doi.org/10.1177/1461444818799523>.
- PASQUALE Frank (2015), *The black box society*, Cambridge, Harvard University Press.
- PELIZZA Annalisa et al. (2021), « Identification as translation: The art of assembling European migration data », *Science, Technology, & Human Values*, 46(3), p. 530-561, DOI : <https://doi.org/10.1177/016224392090443>.
- PELIZZA Annalisa, Stefania MILAN et Yoen LAUSBERG (2021), « Understanding migrants in COVID-19 counting: Rethinking the data-(in)visibility nexus », *Data and Policy*, 3(4), DOI : <https://doi.org/10.1017/dap.2021.19>.
- PÈNE Sophie (2020), « Métiers de la fonction publique : motifs et modèles de transition numérique », *Approches théoriques en information-communication*, 1(1), p. 58-80.

PEUGEOT Valérie (2018), « Données de santé : contours d'une controverse », *L'Économie politique*, 80(4), p. 30-41.

PINK Sarah, Debora LANZENI et Heater HORST (2018), « Data anxieties: Finding trust in everyday digital mess », *Big Data & Society*, 5(1), DOI : <https://doi.org/10.1177/2053951718756685>.

PLANTIN Jean-Christophe, Carl LAGOZE, Paul N. EDWARDS et Christian SANDVIG (2017), « Big data is not about size: When data transform scholarship », dans Clément MABI et Laurence MONNOYER-SMITH (dir.), *Ouvrir, partager, réutiliser. Regards critiques sur les données numériques*, Paris, Éditions de la Maison des sciences de l'homme, DOI : <https://doi.org/10.4000/books.editionsmsmh.9103>.

PRINSLOO Paul (2019), « A social cartography of analytics in education as performative politics », *British Journal of Educational Technology*, 50(6), DOI : <https://doi.org/10.1111/bjet.12872>.

PRINSLOO Paul, Sharon SLADE et Mohammad KHALIL (2022), « The answer is (not only) technological: Considering student data privacy in learning analytics », *British Journal of Educational Technology*, 53(4), DOI : <https://doi.org/10.1111/bjet.13216>.

PRONZATO Riccardo et Annette N. MARKHAM (2023), « Returning to critical pedagogy in a world of datafication », *Convergence*, 29(1), DOI : <https://doi.org/10.1177/13548565221148108>.

PURWANTO Ahmad, Anneke ZUIDERWIJK et Marijn JANSSEN (2020), « Citizen engagement with open government data: A systematic literature review of drivers and inhibitors », *International Journal of Electronic Government Research*, 16(3), p. 1-25, DOI : <https://doi.org/10.4018/IJEGR.2020070101>.

RAINIE Carroll, Tahu KUKUTAI, Maggie WALTER, Oscar Luis FIGUEROA-RODRIGUEZ, Jennifer WALKER et Per AXELSSON (2019), « Issues in open data: Indigenous data sovereignty », dans Tim DAVIES, Stephen B. WALKER, Mor RUBINSTEIN et Fernando PERINI, *The state of open data: Histories and horizons*, Cape Town/Ottawa, African Minds/International Development Research Centre, DOI : <https://doi.org/10.5281/ZENODO.2677801>.

RICHTERICH Annika (2018), *The big data agenda: Data ethics and critical data studies*, Londres, University of Westminster Press.

RIEDER Gernot et Judith SIMON (2016), « Datatrust: Or, the political quest for numerical evidence and the epistemologies of big data », *Big Data & Society*, 3(1), DOI : <https://doi.org/10.1177/2053951716649398>.

ROGAN Kevin (2019), « The universal factory: Data production and platforms », *Enquiry*, 16(2), p. 18-31, DOI : <https://doi.org/10.5840/enquiry20191624>.

ROSALES Andrea et Mireia FERNÁNDEZ ARDÈVOL (2019), « Structural ageism in big data approaches », *Nordicom Review*, 40(1), p. 51-64, DOI : <https://doi.org/10.2478/nor-2019-0013>.

RUBERG Bonnie et Spencer RUELOS (2020), « Data for queer lives: How LGBTQ gender and sexuality identities challenge norms of demographics », *Big Data & Society*, 7(1), DOI : <https://doi.org/10.1177/2053951720933286>.

RUPPERT Evelyn (2015), « Doing the Transparent State: Open government data as performance indicators », dans Richard ROTTENBURG, Salle E. MERRY, Sung-Jonn PARK et Johanna MUGLER (dir.), *A world of indicators: The making of governmental knowledge through quantification*, Cambridge, Cambridge University Press, p. 127-150.

RUPPERT Evelyn (2019), « Different data futures: An experiment in citizen data », *Statistical Journal of the IAOS*, 35(4), p. 633-641.

- SANDER Ina (2020), « What is critical big data literacy and how can it be implemented? », *Internet Policy Review*, 9(2), DOI : <https://doi.org/10.14763/2020.2.1479>.
- SCHMIDT Harald, Dorothy E. ROBERTS et Eneanya D. NWAMAKA (2021), « Rationing, racism and justice: Advancing the debate around “colourblind” COVID-19 ventilator allocation », *Journal of Medical Ethics*, 48, p. 126-130.
- SCHMIDT Harald, Tatjana SEITZ et Sebastian SPIEGL (2021), « Smart city and data protection: Regulatory challenges of urban data spaces », *Digital Society*, 1(1), DOI : <https://doi.org/10.1016/j.iiref.2025.104022>.
- SELWYN Neil, Luci PANGRAZIO et Bronwyn CUMB (2022), « Knowing the (datafied) student: The production of the student subject through school data », *British Journal of Educational Studies*, 70(3), DOI : <https://doi.org/10.1080/00071005.2021.1925085>.
- SILES Ignacio (2023), *Living with algorithms: Agency and user experience*, Cambridge, Polity.
- SILVER David (2004), « Internet/cyberculture/digital Culture/new media/fill-in-the-blank studies », *New Media & Society*, 6(1), DOI : <https://doi.org/10.1177/1461444804039915>.
- SOUTHERTON Claire (2022), « Datafication », dans Laurie A. SCHINTLER et Connie L. MCNEELY (dir.), *Encyclopedia of big data*, Cham, Springer International Publishing, p. 358-361.
- STARK Luke et Anna Lauren HOFFMANN (2019), « Data is the new what? Popular metaphors & professional ethics in emerging data culture », *Journal of Cultural Analytics*, 4(1), DOI : <https://doi.org/10.22148/16.036>.
- STEVENS Marthe, Rik WEHRENS et Antoinette DE BONT (2018), « Conceptualizations of big data and their epistemological claims in healthcare: A discourse analysis », *Big Data and Society*, 5(2), DOI : <https://doi.org/10.1177/2053951718816727>.
- TAFFEL Sy (2021), « Data and oil: Metaphor, materiality and metabolic rifts », *New Media & Society*, 25(5), p. 980-998, DOI : <https://doi.org/10.1177/1461444821991023>.
- TAYLOR Linnet (2017), « What is data justice? The case for connecting digital rights and freedoms globally », *Big Data & Society*, 4(2), DOI : <https://doi.org/10.1177/2053951717736335>.
- THEVIOT Anaïs (2016), « Les data: nouveau trésor des partis politiques ? Croyances, constitutions et usages comparés des données numériques au Parti socialiste et à l'Union pour un mouvement populaire », *Politiques de communication*, 1(6), p. 137-166.
- THEVIOT Anaïs (dir.) (2023), *Gouverner par les données ? Pour une sociologie politique du numérique*, Paris, Presses des Mines.
- THEVIOT Anaïs et Anne BELLON (2024), dossier « Big Data : usages et résistance », *Quaderni*, 112.
- THEVIOT Anaïs et Éric TREILLE (2019), « Les “big data” au travail. Les métiers de la donnée entre expertises professionnelles et effets de croyance », *Politiques de communication*, 12(1), p. 5-20.
- VAN DIJCK José (2014), « Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology », *Surveillance & Society*, 12(2), p. 197-208.
- VAN DIJCK José et Thomas POELL (2016), « Understanding the promises and premises of online health platforms », *Big Data & Society*, 3(1), DOI : <https://doi.org/10.1177/2053951716654173>.
- VENTURINI Tommaso, Amelia ACKER, Jean-Christophe PLANTIN et Tone WALFORD (dir.) (2025), *The SAGE handbook of data and society*, New York, Sage Publications.
- VERDI Ugo (2023a), « L'éthique des données dans les chartes éthiques des collectivités territoriales », *Communication & Organisation*, 64, p. 51-66.

VERDI Ugo (2023b), « Quelle(s) réponse(s) à l'enjeu d'acculturation aux données ? Un état de l'art des caractéristiques de la data literacy », *Revue française des sciences de l'information et de la communication*, 26, DOI : <https://doi.org/10.4000/rfsic.14589>.

WALTER Maggie, Raymond LOVETT, Bobby MAHER, Bhiamie WILLIAMSON, Jacob PREHN, Gawaian BODKIN-ANDREWS et Vanessa LEE (2020), « Indigenous data sovereignty in the era of Big Data and Open Data », *Australian Journal of Social Issues*, 56(2), p. 143-156.

WAMSLEY Dillon et Benjamin CHIN-YEE (2021), « COVID-19, digital health technology and the politics of the unprecedented », *Big Data & Society*, 8(1), DOI : <https://doi.org/10.1177/20539517211019441>.

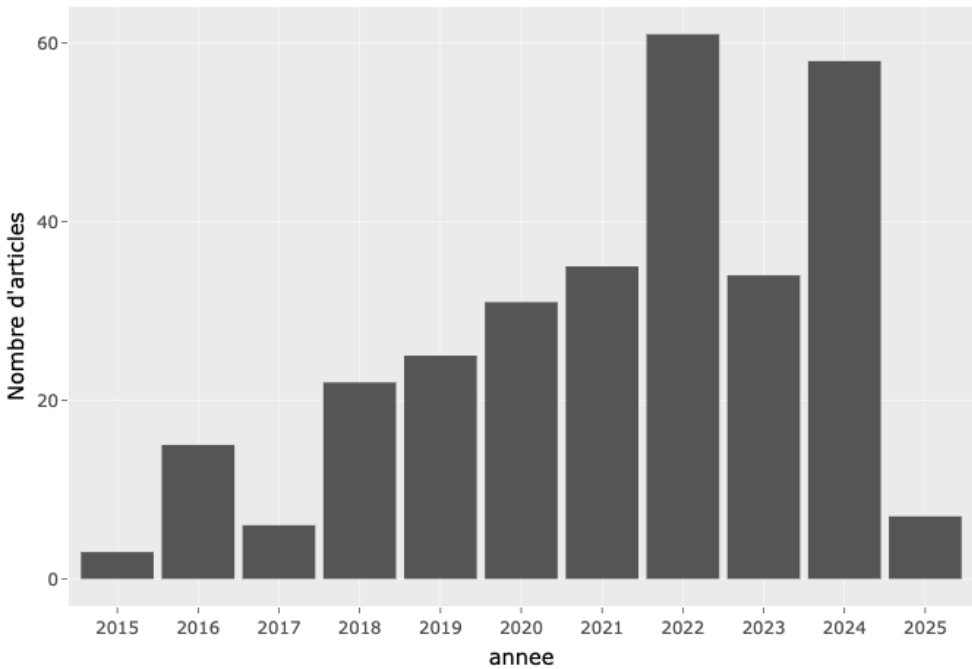
WATSON-DANIELS Jamelle, Yeshimabeit MILNER, Nicole TRIPLETT, Irene HEADEAN, Dominique DAY, Zinzi BAILEY... Lucas MASON-BROWN (2020), *Data for Black Lives COVID-19 movement pulse check and roundtable report*, <https://d4bl.org/reports/41-data-for-black-lives-covid-19-movement-pulse-check-and-roundtable-report>, page consultée le 29 novembre 2025.

YOUNG Alyson Leigh et Anabel QUAN-HAASE (2013), « Privacy protection strategies on Facebook: The Internet privacy paradox revisited », *Information, Communication & Society*, 16(4), DOI : <https://doi.org/10.1080/1369118X.2013.777757>.

ZUBOFF Shoshana (2019), *The age of surveillance capitalism: The fight for a human future at the new frontier of power*, New York, PublicAffairs.

ANNEXES

Nombre de publications extraites de Scopus ayant dans leur titre, leur résumé ou leurs mots-clés les mots *critical data studies* (extraction réalisée le 11 février 2025 à l’aide de R), entre 2015 et 2025



Liste des revues dans lesquelles sont publiés les travaux en *critical data studies*, classées en ordre d’importance, avec le nombre de publications correspondant

Big Data & Society	57
New Media & Society	11
Transforming Communication	11
Information, Communication & Society	10
Internet Policy Review	7
Proceedings of the ACM on Human-Computer Interaction	7
International Journal of Communication	5
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	4

Proceedings of the Association for Information Science and Technology	4
Annals of the American Association of Geographers	3

Source : corpus de publications extrait de Scopus le 11 février 2025, voir l'annexe 1.

NOTES

1. La décennie 2010 est celle qui voit s'organiser des travaux clés permettant d'identifier les contours et spécificités du champ (voir Kitchin, 2022).
2. Pour les références en français, les bases de données interrogées incluaient Scopus, CAIRN, Sophia, Open Edition Freemium, Repères, Érudit ([Mots-clés, titre ou résumé] : « études critiques sur les données » OU « Critical Data Studies » OU « approches critiques des données » OU « sociologie de la quantification » OU « Approches critiques de la quantification » OU « biais algorithmique » OU « études critiques des données » ou « biais algorithmique »). Pour celles en anglais, les bases de données interrogées incluaient Scopus et Sophia ([mots-clés] : « Critical Data Studies » AND [digital]).
3. L'un des avantages de la recherche automatisée dans Scopus est d'avoir fait émerger des travaux sur la transformation de domaines et secteurs professionnels touchés par la *datafication*, dont plusieurs étaient publiés dans des revues propres à ces domaines, donc plus difficiles à repérer manuellement. Ce nouveau portrait du champ a aussi permis de réorganiser partiellement nos thèmes et catégories d'analyse pour en proposer une cartographie. Un inconvénient de cette recherche est qu'elle a omis certains travaux fondateurs qui ne citaient pas explicitement l'expression *critical data studies*, que nos recherches manuelles avaient toutefois permis de repérer. Cette démarche confirme la plus-value d'une stratégie combinant recherche manuelle et semi-automatisée pour ratisser suffisamment large, tout en conservant la main sur l'évaluation de la pertinence des travaux recensés.
4. Précisons que les objectifs de l'analyse bibliométrique se limitent au portrait du champ à partir des éléments évoqués ; ils ne visent ni l'analyse des réseaux de collaboration ni celle de la structure intellectuelle du domaine (similarité des bibliographies). Soulignons, par ailleurs, que même si les données fournies par Scopus sont réputées de bonne qualité, ses limites dans la couverture des sciences humaines et sociales, qui plus est en français, sont reconnues (à titre d'exemple, les chapitres de livres y sont moins bien référencés que les articles de revues).
5. « Datafication refers to the process by which subjects, objects, and practices are transformed into digital data »
6. « CDS understands data as being *contingent* and relational in nature; their generation, processing, and use *contextually* framed (conceptually, technically, socially, economically) and thoroughly infused by politics (personal, institutional, cultural) »
7. D'autres travaux en *critical data studies* s'intéressent à des secteurs comme la culture, le journalisme ou l'agriculture, où la *datafication* est également en cours, ou encore à certains domaines scientifiques caractérisés par un usage important de données, comme la géographie, mais ces travaux restent peu nombreux, d'une part, du moins dans le cadre de notre travail de recension, et leur ancrage dans les études critiques de données est parfois marginal, d'autre part.

RÉSUMÉS

Cet article présente une cartographie du champ des études critiques sur les données (*critical data studies*). Il retrace son émergence au sein de la littérature anglophone et francophone et précise sa perspective sur le plan épistémologique autour de la constitution d'un objet de recherche propre, la « donnée ». Il organise les travaux rassemblés dans ce champ de recherche autour de deux axes thématiques : l'étude des conséquences de la mise en données sur la vie sociale et l'étude des solutions de remplacement à la *datafication*.

This article presents a mapping of critical data studies. It traces the field's emergence in English- and French-language scholarship and clarifies its epistemological perspective by examining the definition of its specific research subject: "data." The article brings together work in this field around two main themes: the study of the social consequences of data and the study of alternatives to datafication.

El artículo presenta una cartografía del campo de estudios críticos sobre los datos (*Critical Data Studies*). Este identifica su aparición en los escritos anglófonos y francófonos y precisa su perspectiva a nivel epistemológico acerca de la creación de un objeto de investigación exclusivo, es decir, el "dato". También organiza los datos recopilados en ese campo de investigación en torno a dos líneas temáticas: El estudio de consecuencias de la configuración de datos sobre la vida social y el estudio de las alternativas a la dataficación.

INDEX

Mots-clés : datafication, mise en données, épistémologie critique, sociotechnique, cartographie

Keywords : datafication, critical epistemology, sociotechnical, mapping

Palabras claves : dataficación, configuración de datos, epistemología crítica, sociotécnica, cartografía

AUTEURS

FLORENCE MILLERAND

Florence Millerand est professeure titulaire au Département de communication sociale et publique de l'Université du Québec à Montréal. Courriel : millerand.florence@uqam.ca

MÉLANIE MILLETTE

Mélanie Millette est professeure titulaire au Département de communication sociale et publique de l'Université du Québec à Montréal. Courriel : millette.melanie@uqam.ca

LUCIE DELIAS

Lucie Delias est maîtresse de conférences en sciences de l'information et de la communication au département info-com de l'Université de Montpellier Paul-Valéry. Courriel : lucie.delias@univ-montp3.fr

Lectures

Yves WINKIN et Jean-Marie CHARPENTIER (2025), *La communication au long cours. Conversations sur les sciences de la communication*

Caen, Éditions C&F, Coll. « Recherche »

Alexandre Eyries

RÉFÉRENCE

Yves WINKIN et Jean-Marie CHARPENTIER (2025), *La communication au long cours. Conversations sur les sciences de la communication*, Caen, Éditions C&F, Coll. « Recherche »

- 1 Yves Winkin est un chercheur internationalement reconnu qui a joué un rôle essentiel de passeur entre la sociologie d'Erving Goffman, l'anthropologie de la communication de Ray Birdwhistell et les sciences de l'information et de la communication à la française. Jean-Marie Charpentier est un consultant en communication ayant exercé longtemps en entreprise et assumé des responsabilités syndicales. Dans ce nouvel ouvrage, Winkin retrace lors de ses cinq entretiens avec Charpentier le fil de son évolution intellectuelle, de sa carrière et de ses réflexions sur la communication dans une perspective anthropologique et sociale. La richesse et le foisonnement de l'ouvrage m'empêchent de prétendre à l'exhaustivité dans cette recension critique et m'amènent à établir des choix pour donner à voir la pertinence de l'ensemble. Dans l'avant-propos intitulé « À l'origine des conversations » (p. 11-13), Charpentier rappelle dans quelles circonstances il a découvert les travaux de Winkin, soit à travers l'ouvrage collectif *La nouvelle communication* qu'il avait dirigé au Seuil en 1981 et qui, issu de ses séjours aux États-Unis, défrichait un champ entre l'individuel et le collectif, le groupe et l'institution :

Yves Winkin était le passeur de ces recherches qu'on a trop vite résumées à l'école de Palo Alto. Pour le jeune communicant que j'étais, ce livre a servi de premier vaccin contre les tentations dominantes, à la fois technicistes et promotionnelles, du métier. Il y avait autre chose dans la communication et cette autre chose avait à voir avec la relation plutôt qu'avec la transmission (p. 11).

- 2 Les échanges d'abord distanciels vont se poursuivre jusqu'à la première rencontre en face à face qui en appellera d'autres et qui, entre la pandémie de COVID-19 et les nouveaux livres publiés par Winkin, fournira la matière d'autres interactions autour de sa trajectoire professionnelle riche et de son cheminement intellectuel remarquable. Le présent livre en reprend le principe structurant, comme l'écrit Charpentier :

[...] en retraçant ses chemins, il me parle avec simplicité et profondeur d'interaction, d'institution, de communication et aussi d'enchantement. Une notion qu'il affectionne et qui, comme la communication, se révèle ambivalente et féconde. Toujours, avec cette référence-clé de l'anthropologie : le terrain. Je retrouve dans ces conversations la trace et les rebonds de cette communication à la fois sociale et culturelle qui m'a été si souvent utile en situation, tant il est vrai que la communication a d'abord à faire avec les relations qu'elle permet, bien avant toute information ou contenu (p. 13).

- 3 Dans la première conversation, « La nouvelle communication » (p. 15-40), Charpentier rappelle en substance une phrase de l'anthropologue Margaret Mead selon laquelle une communication peut être aussi courte qu'une interaction et aussi longue qu'une génération, ce qui illustre deux dimensions de la communication : l'infiniment petit d'une simple interaction et l'infiniment grand de ce qu'on pourrait appeler une « mémoire sociale ». Winkin revient sur la notion d'anthropologie de la communication qu'il a forgée et rappelle sa volonté d'insister sur ce qu'on appelle en anthropologie la « part ethnographique », c'est-à-dire

le contact immédiat par observation avec ce qu'il se passe autour de soi. Vous vous entretenez avec telle ou telle personne, vous n'êtes pas en train de lui imposer un questionnaire. Il n'y a pas de filtre, pas d'écran. Vous êtes en immersion partielle ou totale. Tous vos sens sont en éveil. Du point de vue de la communication, il n'y a pas seulement la part verbale à récolter et ensuite à analyser. Vous allez aussi essayer de prêter attention au non-verbal, à ce qu'on appelle le « multicanal » (p. 17-18).

- 4 Winkin insiste sur la dimension d'insertion dans un écosystème de communication dont on est partie prenante en rappelant la citation de Birdwhistell, qui avait coutume de dire qu'on ne communiquait pas, mais qu'on participait à la communication, qu'on était plongé dans milieu à part entière :

[...] en anthropologie, vous êtes dans le flot. Vous vous laissez glisser avec le flot et, de temps en temps, vous essayez de vous hisser un peu sur le bord pour regarder ce qu'il se passe. Mais vous gardez quand même les pieds dans l'eau. Vous ne pouvez jamais être comme quelqu'un qui regarde de l'extérieur de l'aquarium nager les poissons. Vous êtes avec les poissons. En anthropologie, il y a toujours immersion (p. 18).

- 5 Profondément inspiré par la pensée de la communication de l'anthropologue américain Birdwhistell, dont il a suivi les cours brillants et disruptifs à l'Université de Pennsylvanie, Winkin rappelle que ce dernier avait coutume de dire que

la culture, vous pouvez l'aborder par ses structures et vous pouvez l'aborder en tant qu'elle est un accomplissement de la culture. La culture se jouant et se produisant justement dans et par les interactions. Interactions encore une fois entre les vivants et pas seulement entre les humains (p. 29).

- 6 S'il doit à Pierre Bourdieu sa première rencontre avec Erving Goffman, dont il deviendra un spécialiste internationalement reconnu, Winkin rappelle que celui-ci pouvait se montrer dans certaines circonstances extrêmement accueillant, même s'il était globalement un mauvais enseignant colérique et peu patient :

[...] quand on est admis à son séminaire, ses étudiants ont un accès direct à lui et à sa documentation personnelle. Il y a d'une part son bureau avec sa secrétaire et d'autre part un autre bureau avec de grands tiroirs métalliques et des classeurs suspendus. Ses étudiants deviennent ses jeunes collègues et il accepte que l'on ait accès à tous les tiroirs rassemblant sa documentation (p. 30).

- 7 Dans la deuxième conversation, « Anthropologie et enchantement » (p. 41-58), Winkin écrit :

[...] la connexion de Pierre Bourdieu avec les États-Unis est forte à tous égards. C'est relativement peu connu. Il n'appartient en tout cas pas à cette génération d'universitaires français anti-américains des années 1960-1970. Bourdieu était à l'Institut d'études avancées de Princeton au début des années 1970. Quand vous lisez « Esquisse d'une théorie de la pratique », il fait référence aux travaux de Harold Garfinkel en ethnométhodologie, alors même qu'en France, on n'en parlait pas encore. Je dirais qu'il était à l'avant-garde de la connaissance en France de ce qui était en train de se faire aux États-Unis. Et il y avait chez lui un attrait réel pour la recherche tant en Amérique du Nord qu'en Amérique du Sud. Je pense au Brésil notamment (p. 45).

- 8 Le chercheur rappelle qu'il a dû au soutien et à l'appui de professeurs de l'Université de Liège, notamment Jacques Dubois, son directeur de thèse, de pouvoir créer au sein du Département de communication un cursus complet d'« anthropologie de la communication » de la première à la quatrième année, spécialité qu'il va avoir l'occasion de pratiquer concrètement à l'occasion de multiples déplacements :

[...] c'est très curieux. Ce sont pour commencer des terrains un peu opportunistes. Je vais faire flèche de tout bois. Je vais avoir l'occasion de pas mal voyager, de faire des randonnées à l'étranger avec un voyageur comme Terre d'aventure, par exemple. Je vais aussi accepter des missions d'enseignement dans différents pays (p. 49).

- 9 Des voyages naît l'anthropologie de l'enchantement qui souligne, à travers la massification de l'expérience touristique, le besoin de l'homme de suspendre un temps sa propre incrédulité pour croire à la mise en scène qui lui est proposée, une utopie concrète en somme. La troisième conversation, « Communication et culture » (p. 59-82), donne l'occasion à Winkin de préciser sa vision de la communication qui rend toujours actuelle et dynamique la culture : « [...] la culture est performée par le fait même qu'elle est accomplie dans l'ici et maintenant de l'interaction communicationnelle. Elle est réalisée dans le sens fort du terme. Elle devient quelque chose de concret, de presque tangible » (p. 62). Ainsi, la culture ne reste pas seulement un ensemble de codes secrets et compliqués, comme le disait Edward Sapir. À un moment donné, dans l'interaction, la culture émerge et se donne à voir dans les comportements mêmes des membres de cette culture. Lucide par rapport à la vision dégradée que la plupart des gens ont de la communication, Winkin rappelle fort à propos que

pour nombre de sciences sociales en dehors des sciences de l'information et de la communication, le mot communication reste un mot un peu suspect qui connote la gestion, les affaires, les relations publiques. C'est vu non pas comme un concept, mais comme une action et de plus une action vénale, une action commerciale. « Il est dans la communication », ça veut dire il est dans le bluff, il est dans le faire semblant (p. 90).

- 10 Dans l'après-propos, « Passer le témoin » (p. 127-130), le chercheur note avec justesse qu'il
- faudrait un jour analyser l'étonnante alchimie interactionnelle et cognitive de l'entretien à distance. Les pages qu'on vient de lire sont des « bébés zoom » et elles ont produit des choses étonnantes pour moi : remontées de souvenirs à la surface, tressage de fils rouges encore jamais perçus, aveux très personnels (p. 127).
- 11 À l'occasion de cet après-propos, qui clôt remarquablement cette série d'entretiens avec Charpentier, Winkin donne une merveilleuse définition de cette « communication au long cours » qui irrigue toutes les conversations qui composent ce livre :
- [...] quand un enseignant-chercheur senior parle de ses années de formation, avec des professeurs depuis longtemps disparus, dont il se souvient qu'ils lui parlaient de leurs propres enseignants, il donne le moyen à des plus jeunes de remonter jusqu'au début du xx^e siècle ou presque. Je me souviens personnellement de Birdwhistell qui nous parlait dans les années 1970 de Margaret Mead qui lui parlait dans les années 1950 de Franz Boas son maître des années 1950 (p. 128).
- 12 Riche et foisonnant, ce livre de Winkin constitue un témoignage inestimable sur l'institutionnalisation et la légitimation progressives des sciences de l'information et de la communication auxquelles il a activement participé en étant à la fois maître d'œuvre et passeur. C'est un ouvrage remarquable qui intéressera à coup sûr les enseignants-chercheurs, formateurs et étudiants passionnés par la communication et dont je recommande vivement la lecture urgente.
-

AUTEURS

ALEXANDRE EYRIES

Alexandre Eyries, Université Catholique de l'Ouest, Laboratoire Centre Humanités et Sociétés (CHUS). Courriel : alex.eyries@yahoo.fr