

Volume 13, numéro 1 / mars 1984 / ISSN 0315-9930

Numéro Spécial: Sciences de l'information

ARGGUS
ARGGUS
ARGGUS

Corporation des bibliothécaires professionnels du Québec

Corporation of Professional Librarians of Québec

ARGUS

Comité de rédaction / Editorial Committee

Maryse Pierrard-Biniak, présidente
Diane Baillargeon
Paulette Bernhard
Denise Diamond
Suzanne Ledoux
Claudine Lussier
Sylvie Painchaud
Suzanne Robitaille

Correcteurs / Correctors

Maryse Pierrard-Biniak
Michel Thériault

Impression / Printing

Atelier des Sourds Montréal (1978) inc.

Traductrice / Translator

Mariene O'Hare-Giguère

Publicité / Advertising

Agence Marsy
440, boulevard Dorchester ouest
Suite 1101
Montréal, Qué. H2Z 1V7
(514) 671-1757

Dépôt légal / Legal deposit

Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ARGUS

Argus est une revue trimestrielle publiée par la Corporation des bibliothécaires professionnels du Québec.

Elle vise à l'information et à l'éducation de la profession.

La rédaction laisse aux auteurs l'entière responsabilité de leurs textes.

L'abonnement annuel est de 25\$ (6,50\$ le numéro) au Québec, 30\$ (7,50\$ le numéro) au Canada et 30\$ U.S. (7,50\$ U.S. le numéro) pour l'extérieur du Canada. Les étudiants bénéficient d'un prix spécial de 15\$ pour l'abonnement annuel.

Toute demande concernant les numéros manquants doit être envoyée, au plus tard un mois suivant la date de parution, à l'adresse suivante:

Argus
Secrétariat de la CBPLQ
360, rue Le Moyne
Montréal, Québec H2Y 1Y3

Toute reproduction des articles, en totalité ou en partie, doit être faite avec la mention de la source.

Les articles de la revue sont indexés dans le *Bulletin signalétique, section 101; Sciences de l'information - documentation, Information Science Abstracts, Library Information Science Abstracts (LISA), Library Literature* et dans *Points de repères*.

Argus is a quarterly magazine sponsored by the Corporation of Professional Librarians of Québec.

Its aim is to publish original papers for the information and education of the profession.

Articles are the entire responsibility of the authors.

The yearly subscription is 25\$ (6,50\$ an issue) in Québec, 30\$ (7,50\$ an issue) in Canada and 30\$ U.S. (7,50\$ U.S. an issue) outside Canada. Students' special rate is 15\$ for the yearly subscription.

Any request concerning missing issues should be sent, no later than a month after date of publication, to the following address:

Argus
Secrétariat of the CPLQ
360 Le Moyne Street
Montréal, Québec, H2Y 1Y3

Articles may be reproduced, in whole or in part with proper credit to their source.

Articles are indexed in *Bulletin Signalétique, section 101; Sciences de l'information - documentation, Information Science Abstracts, Library Information Science Abstracts (LISA), Library Literature* and in *Points de repères*.

VOLUME 13, NUMÉRO 1/MARS 1984/ISSN 0315-9930

ARGUS

Corporation des bibliothécaires professionnels du Québec
Corporation of Professional Librarians of Québec

La lettre d'Argus, *page 3*

Les sciences de l'information: aspects théoriques et interdisciplinaires

Par Jean Tague

page 5

La recherche en science de l'information dans les différents
départements de l'Université de Montréal

Par Suzanne Bertrand-Gastaldy

page 11

Modelling Problems in Bibliometrics

By John J. Hubert

page 19

Recherche sur l'interface système-usager des services de télé-référence

Par Gilles Deschatelets

page 25

Supplément encarté au centre:
Index annuel, volume 12 (1983)



Logiciels pour Micro-ordinateurs **APPLE** (T.M.)

Bientôt disponibles pour autres Micro-ordinateurs

- 1) **Système de prêt en direct**
 - capacité: 65 000 titres, 11 400 abonnés utilisent un crayon lecteur et des codes zébrés.
 - statistiques, inventaires, et plus...
- 2) Système pour acquisitions
- 3) Système pour les jeux de fiches et les étiquettes
- 4) Système de catalogage
- 5) Système de contrôle des publications en série

Bibliofiches (514) 336-4340

Tous les clients des régions (418), (819) et (613) peuvent nous rejoindre directement sans frais en composant: 1-800-361-9211.

Le 8 avril 1983, l'École de bibliothéconomie de l'Université de Montréal organisait un Colloque sur le thème: «Sciences de l'information: recherche et interdisciplinarité». S'inscrivant dans le cadre du dixième anniversaire de la Faculté des arts et des sciences, il n'aurait pu voir le jour sans la collaboration active de Jean Tague. La qualité des conférenciers issus de plusieurs disciplines et de divers milieux de recherche a grandement contribué au succès de cette activité.

Pour le bénéfice de nos lecteurs, nous donnons ci-dessous la liste des interventions:

INTRODUCTION

1. Jean Tague, professeur titulaire et maintenant doyenne à la School of Library and Information Science, University of Western Ontario.
«Les sciences de l'information: aspects théoriques et interdisciplinaires».
2. Suzanne Bertrand-Gastaldy, professeur adjoint, École de bibliothéconomie, Université de Montréal.
«La recherche en sciences de l'information dans les différents départements de l'Université de Montréal».

TRAITEMENT AUTOMATIQUE DES TEXTES

3. Gérard Salton, professeur titulaire, Département d'informatique, Cornell University.
«Extended Boolean Information Retrieval».
4. Richard Kittredge, professeur titulaire, Département de linguistique, Université de Montréal.
«Structures d'information dans certains sous-langages techniques».

BIBLIOMÉTRIE ET SCIENTOMÉTRIE

5. John Hubert, professeur agrégé, Département de mathématiques et statistiques, University of Guelph.
«Modelling Problems in Bibliometrics».
6. Yakov Rabkin, professeur agrégé, Institut d'histoire et de socio-politique des sciences, Université de Montréal.
«L'analyse des citations dans le cadre de la recherche sur la recherche».

COMMUNICATION

7. Gilles Deschatelets, chef de la division des collections scientifiques, Université Laval.
«Recherche sur l'interface système-usager des services de téléconférence».
8. Michel Cartier, directeur, Laboratoire de télématique, Université du Québec à Montréal.
«Télématique et médiatique: le projet Agora».
9. Jean-Eudes Bériault, agent de réseau, Bibliothèque nationale du Canada.
«iNet et les bibliothèques canadiennes».

La publication de quatre des conférences (les première, deuxième, cinquième et septième) permettra aux lecteurs d'*Argus*, qu'ils aient ou non participé au Colloque, d'avoir un bon aperçu des réflexions qui se poursuivent dans le domaine de la Science de l'information et de prendre la mesure de la diversité et de la richesse des activités menées actuellement.

Pour les textes non présentés dans ce numéro, nous renvoyons aux publications récentes de leurs auteurs.

Bonne lecture à tous.

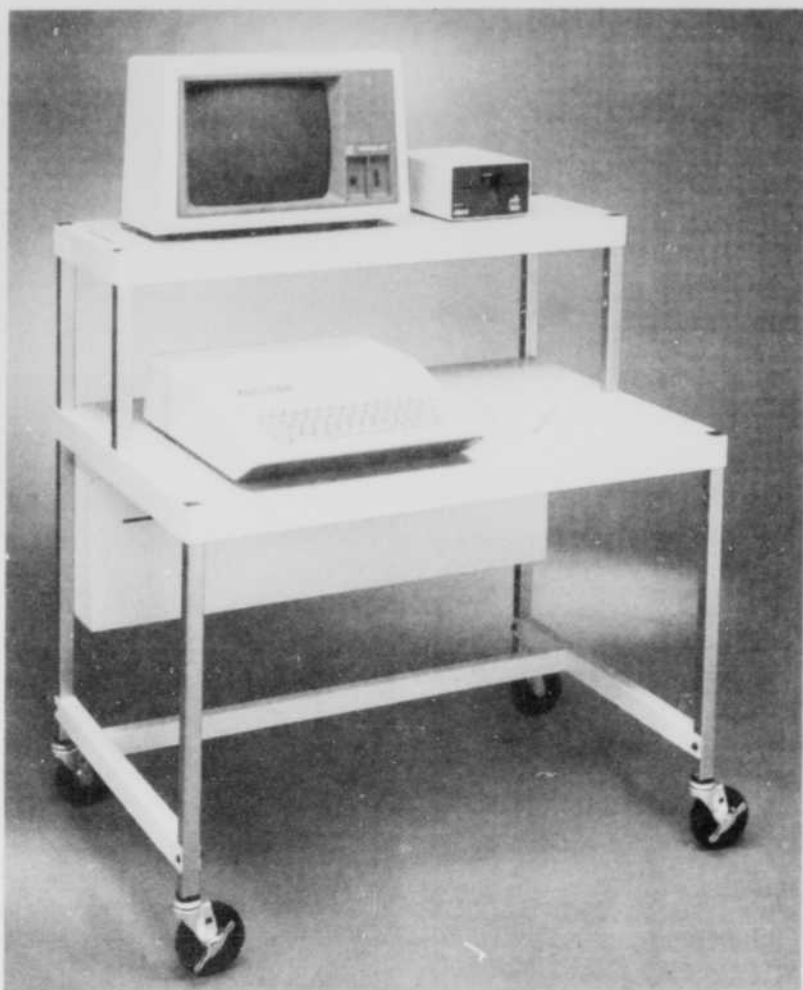


Table à micro-ordinateur

- Fabriquée de métal robuste
- Dessus en métal (36" L x 24" P) 91,8 cm x 61,2 cm, hauteur ajustable (27") 68,85 cm ou moins; émail cuit, couleur ivoire.
- Tablette ajustable en métal (36" L x 12" P) 91,8 cm x 30,6 cm hauteur ajustable (39") 99,45 cm ou moins; émail cuit, couleur ivoire.
- 3 prises de courant électrique avec corde de rallonge (20') 6 mètres approx.
- 4 roulettes à billes (4") 10,2 cm (2 munies de freins).
- Corbeille à l'arrière du dessus pour recevoir les cordes de l'ordinateur et des périphériques.

No 08-100 Prix sur demande

Les sciences de l'information: aspects théoriques et interdisciplinaires

Par Jean Tague*

L'auteur s'interroge sur le contenu du concept de «sciences de l'information» et trouve que cette discipline, encore neuve et en évolution constante, n'a pas encore pu intégrer ses diverses facettes dans un tout cohérent exprimé par une théorie globale: l'auto-examen n'a pas encore abouti à une auto-définition claire. L'auteur, bien sûr, essaie de cerner cette théorie en analysant les quatre composantes de la discipline des sciences de l'information: le repérage de l'information, la bibliométrie, la mesure de l'information et les aspects technologiques et sociaux de la communication.

The author questions the concept of information science and proves that this discipline, still new and evolving, has not yet integrated its diverse aspects into a coherent global theory. Auto-examination has not yet led to auto-definition. The author, of course, attempts to define this theory by analysing four components of information science: information retrieval, bibliometrics, the measurement of information and the technical and social aspects of communication.

La science de l'information est-elle vraiment une science? Ou bien est-elle simplement un mélange de concepts, d'approches, de théories, empruntés aux autres disciplines?

S. Bertrand-Gastaldy et D. Reicher décrivent le développement de la science de l'information comme une progression de la bibliothéconomie par la voie de la documentation:

«De la classification du livre au traitement de l'information transcrite sur d'autres supports, puis à l'analyse approfondie des données enregistrées en mémoire d'ordinateur, on est passé à une réflexion théorique sur le phénomène global de l'information et du transfert des connaissances. On en est venu progressivement à réfléchir sur les difficultés culturelles, psychologiques et philosophiques à surmonter pour rendre disponible tout le savoir humain.»

Mais, pour être une science, une étude doit être plus qu'une utilisation de technologies diverses. Donc, pour

répondre à la question posée, il faut se demander d'abord ce qu'est une science.

On peut distinguer quatre aspects dans toute science:

1. La description des phénomènes naturels qu'on étudie ou les problèmes qu'on veut résoudre;
2. La définition des concepts importants de la discipline;
3. La formulation des hypothèses et ensuite des lois;
4. L'intégration des lois dans une théorie ou un modèle souvent mathématique avec lequel on peut faire des prévisions.

Quel est l'état actuel de la science de l'information par rapport à ces aspects? Il y a des phénomènes: les média, les documents, les index et les catalogues, les utilisateurs, les réseaux de communication. Il y a des définitions et des hypothèses, présentées par beaucoup d'auteurs, comme les lois de Zipf et de Bradford. Mais la discipline n'est pas encore parvenue à intégrer ces lois dans une théorie globale. Dans cet article, nous allons présenter la science de l'information d'une manière un peu personnelle. Mais, de plus, nous avons l'intention de procéder à une

intégration préliminaire de la discipline.

Quel est le phénomène central qu'étudient les scientifiques de l'information? Ce n'est pas l'information en son sens général, mais plutôt la description et l'utilisation des documents, c'est-à-dire des messages fixés sur des supports. Un document est une représentation de la pensée humaine. Il peut être oral, visuel, écrit, imprimé, ou enregistré en mémoire d'ordinateur. Un texte est simplement un document qui utilise le langage. Comme Long² le note, un texte a un aspect strictement linguistique et aussi un aspect graphique. Un texte est fait de symboles discrets. D'ailleurs, c'est un arrangement ou une structuration des symboles. Par exemple, ce peut être une série ou une table de ces symboles. Un symbole est la plus petite marque sémantique, c'est-à-dire qu'un symbole est une marque qui a un sens. Ce peut être un mot ou un lexème ou un code binaire. L'ensemble des symboles produit le sens ou la sémantique du texte. Cette sémantique signifie vraiment deux choses qui agissent ensemble:

1. La correspondance entre ce texte et les autres textes

* L'auteur est professeur titulaire à la School of Library and Information Science, University of Western Ontario, London, Ontario et en deviendra la doyenne le 1^{er} juillet 1984.

2. La correspondance entre le texte et un phénomène réel quelconque.

La science de l'information vise à préciser la nature de ces correspondances.

Définition de l'information

On peut définir l'information de plusieurs façons. Quelques auteurs la définissent comme une opération ou un processus, la communication d'une personne à une autre ou la transmission d'un message ou le changement d'une image mentale. Selon Shannon, l'information est la réduction de l'incertitude d'un récepteur à propos du contenu d'un message transmis d'une source par un canal. D'autres définissent l'information comme un produit, comme un ensemble de données ou de chiffres, comme la réponse à une question, comme les résultats d'une expérience, comme le texte lui-même. Finalement, pour d'autres, l'information est une caractéristique ou une qualité. C'est la rareté d'un message reçu, la complexité d'un logiciel, la grandeur d'une collection, l'utilité de la réponse à une question ou des conséquences d'une décision.

Le philosophe Karl Popper situe dans un contexte les discussions de la définition de l'information dans son livre *Objective Knowledge*.³ Il distingue trois mondes: le monde des objets physiques, le monde des états subjectifs de l'esprit, et le monde de la connaissance objective. Le troisième monde représente ici les contenus logiques des livres, des revues, des journaux. Il comprend les théories, les problèmes, les arguments, les conjectures, les mythes, les paradigmes. Il est basé sur le document, le support physique, mais il est vraiment un produit de l'esprit. Quelques spécialistes ont douté de l'existence d'un troisième monde distinct du premier et du deuxième. Mais comme Brookes⁴ le remarque, si l'on imagine un monde sans humanité, on peut encore imaginer l'existence de la connaissance. Si des textes restent, on n'a besoin que d'un traducteur pour révéler cette connaissance.

Quelle est la relation entre la connaissance objective et l'information? Comme nous l'avons déjà indiqué, on peut regarder l'information comme un objet du premier monde (les documents eux-mêmes) ou un objet du deuxième monde (une image mentale que les documents transmettent). Mais si l'on accepte l'existence du troisième monde, il me semble que l'information

est vraiment un processus: l'action du troisième monde sur le deuxième monde. Quand on lit ou quand on passe en revue des documents, à mesure qu'on comprend ce qu'on lit, on reçoit ou on reconnaît des objets du troisième monde. Si ces objets du troisième monde sont tous nouveaux pour le lecteur, ils modifient son état mental. Cette modification est l'information.

Spécialités de la science de l'information

Si on lit régulièrement les revues de la discipline, comme *Journal of Documentation*, *Information Processing and Management*, *Journal of the American Society for Information Science*, *Documentaliste*, *Sciences de l'information*, il est clair que la recherche en science de l'information se divise en quatre spécialités:

1. Le repérage de l'information
2. La bibliométrie
3. La mesure de l'information
4. Les aspects technologiques et sociaux de la communication.

Dans une certaine mesure, les chercheurs de chaque sous-champ s'occupent du même phénomène — l'information — mais avec des différences. Chacun est caractérisé par ses propres points de vue, ses propres concepts et théories, ses propres problèmes. Donc, dans la suite de cet article, nous allons examiner tour à tour les recherches et les problèmes qui sont couramment à la frontière de chaque spécialité de la science de l'information.

Le repérage de l'information

L'ordinateur provoque une révolution dans le repérage de l'information documentaire et généralement dans le traitement du texte. Mais c'est une révolution du premier monde, pas vraiment du troisième monde. Les problèmes sont arrivés avant l'ordinateur. Les plus importants sont les suivants:

1. Les structures et les méthodes optimales pour trouver la réponse à une question parmi un ensemble de textes.
2. Le moyen optimal de représenter les textes dans les index.
3. L'effet de la grandeur sur les systèmes de repérage de l'information.
4. Les meilleurs modèles conceptuels de systèmes de repérage de l'information.

Le travail de Salton⁵ est très connu dans ce domaine. Il a développé un système expérimental de repérage de

l'information (le système SMART) basé sur un modèle «espace-vecteurs». Beaucoup d'autres chercheurs suivent son exemple et ces recherches aboutissent aux méthodes probabilistes de repérage, dans lesquelles on essaie de déterminer la probabilité de pertinence d'un document en réponse à une question en fonction de caractéristiques quantitatives des documents et des mots-clés. On peut trouver une description de ces méthodes et d'autres systèmes expérimentaux de repérage dans les comptes rendus du Congrès international sur la recherche et le développement du repérage de l'information⁶ et dans la monographie bien connue de Van Rijksbergen.⁷

Les systèmes commerciaux en mode direct utilisent pour la plupart la logique booléenne comme stratégie de recherche. Mais, comme le soulignent les disciples du repérage probabiliste, les stratégies booléennes, posent des problèmes. Il semble qu'il ne soit pas possible d'atteindre en même temps des taux élevés de rappel et de précision. Si l'on formule une stratégie plus étroite le taux de rappel diminue et le taux de précision augmente. Si l'on utilise une stratégie plus large, le taux de rappel augmente et le taux de précision diminue.

Quelques chercheurs s'intéressent aux améliorations pouvant être apportées aux systèmes de repérage en mode direct. Deschatelets⁸ étudie le rôle de l'intermédiaire et les négociations entre l'utilisateur et les intermédiaires. Il trouve que la qualité de la recherche est meilleure si les utilisateurs interrogent eux-mêmes ou terminal les bases de données bibliographiques.

D'autres approches font appel à une analyse linguistique profonde de la structure des représentations des documents et des questions dans le système. C'est à ce niveau que les intérêts des chercheurs de la science de l'information approchent ceux des linguistes et des informaticiens qui travaillent dans le domaine de l'intelligence artificielle ou de la traduction automatique. Dans chaque cas, il s'agit de procéder à une analyse syntaxique et sémantique pour mieux comprendre le langage naturel.

Dans beaucoup de systèmes question-réponse propres à l'intelligence artificielle, la grammaire des cas de Fillmore⁹ est employée comme base pour cette analyse syntaxique-sémantique. Mais son utilité pour les systèmes de repérage de l'information n'est pas encore évidente. Le problème, comme le montre la recherche d'une

de nos étudiantes¹⁰ est que l'attribution de cas à des noms qui sont liés à un verbe n'est pas identique d'un analyste à un autre. Le même problème se produit quand on ajoute les relations à des mots-clés comme dans l'indexation relationnelle de Farradane¹¹ ou le système Syntol.¹²

Les réseaux sémantiques sont des formalisations que beaucoup d'informaticiens utilisent pour représenter les relations entre les concepts.¹³ Dans une représentation graphique d'un réseau, les sommets sont les mots ou les concepts et les arcs sont les relations entre les mots ou les concepts. Les sommets et les arcs peuvent tous les deux porter des étiquettes. Cela signifie qu'il y a des sommets et des arcs de différentes sortes. Par exemple, les relations peuvent être genre-espèce, partie-totalité, possesseur-possession, etc. Le tableau 1 donne un exemple de réseau sémantique.

Tableau 1



Les problèmes de la traduction automatique et du repérage de l'information se ressemblent parce que dans les deux cas on doit comprendre le sens du texte. Donc les études des chercheurs dans ce domaine sont valables pour les scientifiques de l'information. Le travail de Wilkes¹⁴ est particulièrement pertinent sur cette question. Il propose une sémantique préférentielle pour distinguer les homographes, laquelle est basée sur un réseau sémantique dans lequel les sommets représentent les primitifs sémantiques. On choisit le sens d'un mot par le contexte. Cependant du point de vue du repérage de l'information une telle analyse est très ardue et coûte très cher.

Actuellement, l'objectif le plus important pour les systèmes de repérage de l'information est de développer des modèles conceptuels de bases de données bibliographiques et les processus de ces systèmes. Les informaticiens ont des modèles conceptuels pour les SGBD (systèmes de gestion des banques de données), c'est-à-dire les systèmes qui contiennent les informations factuelles ou numériques. Crawford¹⁵

essaie d'adapter le modèle relationnel de ce genre aux bases de données bibliographiques. Mais à mon avis, cette adaptation est artificielle. Le concept d'un fichier « plat » est trop simple pour représenter les relations diverses des éléments des bases bibliographiques. On a besoin de modèles naturels, propres à cette espèce d'information. Forest¹⁶ présente un modèle intéressant d'indexation et de repérage. Il faut le développer.

De plus, il y a des bases textuelles, c'est-à-dire des bases en langage naturel. Comment trouve-t-on l'information ici? Chou¹⁷ fait des recherches au sujet de la lecture, un aspect auquel s'intéressent aussi les éducateurs. Il propose un modèle du traitement de l'information écrite par l'humain, un modèle qui intègre les recherches récentes en intelligence artificielle et en sémiologie. Il développe ce modèle pour la consultation, mais il y a bien sûr les autres objectifs de la lecture, comme l'éducation générale et le plaisir. Il faut les distinguer et peut-être développer d'autres modèles qui comprennent les relations entre ces objectifs.

La bibliométrie

La bibliométrie est l'étude des aspects quantitatifs des textes et de tous les comportements qui s'y rattachent. Les lois de Zipf, Lotka, et Bradford sont assez bien connues, mais il n'est pas certain qu'elles soient vraiment des lois, dans le sens que l'évidence empirique n'est pas toujours convaincante. Pour une revue de ces lois et d'autres sujets qui intéressent les bibliométriciens, on peut consulter la livraison de la revue *Library Trends* de l'été 1981.

Zipf¹⁸ remarqua une relation empirique entre le rang d'un mot dans un long texte et sa fréquence. Le produit des deux valeurs est une constante. Plus tard, Mandelbrot¹⁹ proposa une forme plus générale de la loi:

$$f(x) = \frac{c}{(b+x)^d}$$

Il y a deux interprétations de cette formule: l'interprétation grandeur-fréquence ou l'interprétation rang-fréquence. Dans la première interprétation, x représente le nombre d'occurrences d'un mot et $f(x)$ représente la fréquence de mots de ce nombre d'occurrences. Dans la seconde interprétation, x représente le rang d'un mot lorsque les mots sont mis en ordre du plus fréquent au moins fréquent et $f(x)$ représente la fréquence de mots de ce rang. Cette distribution

est une sorte de loi de rendement décroissant qui s'applique à beaucoup d'autres phénomènes comme la grandeur des villes et le nombre d'espèces que comprend un genre. Pour les mots cette relation n'est pas forte et beaucoup de chercheurs, comme le décrit Hubert,²⁰ cherchent une meilleure formule. L'application aux vedettes-matière ou aux mots-clés est utile dans la simulation des systèmes de repérage de l'information, comme l'ont constaté Tague et Nelson.²¹

La loi de Lotka²² est vraiment une spécialisation et une autre application de la formule de Mandelbrot, ou $f(x)$ est le nombre d'auteurs qui ont x publications, et a est le nombre d'auteurs qui ont une seule publication, et $b = 0$, $c = 2$.

La loi de Bradford²³ concerne le rendement de revues ou d'autres sources de publications. C'est aussi une loi du rendement décroissant, mais la forme est différente de celle de Zipf. La forme la mieux connue est:

$$F(r) = a + b \log(r)$$

où r est le rang de chaque revue déterminé par la quantité d'articles publiés sur un sujet donné. Le rang 1 est attribué à la revue qui publie le plus sur le sujet en question; le rang 2 à celle dont le nombre de publications vient immédiatement après et ainsi de suite. $F(r)$ est le nombre cumulé d'articles publiés par les r revues, tandis que a et b sont des paramètres qui dépendent du domaine. La correspondance entre la loi de Bradford et les données empiriques est plus forte dans le cas de la loi de Zipf, mais cette correspondance est due en partie au fait que la loi de Bradford est une distribution cumulative, alors que celle de Zipf est non cumulative. La cumulation a l'effet d'enlever les perturbations.

Le champ le plus actif de la bibliométrie est l'analyse des citations. Price²⁴ et Garfield²⁵ ont été les pionniers dans ce domaine; le second l'applique au repérage de l'information, dans les index des citations de l'Institute for Scientific Information, le premier en développe la théorie. Actuellement, les chercheurs visent à déterminer la vraie signification d'une citation; il s'agit de déterminer pourquoi les auteurs citent et s'il y a une relation entre l'importance ou la qualité d'un ouvrage et sa fréquence de citation. Le travail le plus innovateur est celui de Small,²⁶ qui essaie de déterminer la structure d'une discipline à partir d'une analyse des citations

et co-citations (la fréquence avec laquelle deux publications sont citées ensemble).

D'autres phénomènes qu'étudient les bibliométriciens sont ceux de la croissance et du vieillissement de la littérature et de l'information que contient la littérature. La relation entre ces deux processus fait l'objet d'un débat. Quelques auteurs, comme Line et Sandison,²⁷ pensent que le vieillissement de la littérature est vraiment une illusion, il ne serait que l'effet de la croissance de la masse documentaire. D'autres, comme Griffiths,²⁸ pensent que le vieillissement est réel. La difficulté est de mesurer à la fois la croissance et le vieillissement.

En bref, la bibliométrie pose beaucoup de questions intéressantes, mais les bibliométriciens cherchent encore les bonnes réponses.

La mesure de l'information

Comment mesurer l'information? Pour répondre à cette question, presque tout le monde pense à Shannon.²⁹ La mesure d'information de Shannon est basée sur les probabilités. Elle est dérivée d'un modèle de télécommunication qui comprend une source, un canal, un destinataire et un message qui est transmis de la source au destinataire. L'information d'un message varie en raison inverse du logarithme de sa probabilité. Donc l'information moyenne d'une source est ce qu'on appelle en thermodynamique l'entropie H:

$$H = \sum_i p_i \log p_i$$

où p est la probabilité du i-ème message. Quelquefois on appelle H l'incertitude de la source.

L'entropie est vraiment une mesure de la rareté d'un message, et non du contenu informationnel. D'autres, comme Bar-Hillel³⁰ et Shreider³¹ suggèrent les mesures dites sémantiques. Cependant, ce sont vraiment des mesures logiques de l'information. Pour les employer il faut traduire le langage naturel en un langage formel basé sur la logique symbolique. Hoffman³² propose une mesure de l'information d'un document qui utilise une représentation du document comme réseau d'unités informationnelles (des faits, des chiffres) dans lequel les arcs signifient les relations entre ces unités. Il nous semble que le développement d'une mesure de l'information sémantique doit atteindre à une compréhension plus claire et profonde de la relation entre le texte et les états mentaux

du lecteur. Il est possible qu'une telle mesure soit une mesure ordinaire, basée sur l'ordre des textes quant à leur efficacité informationnelle au lieu d'une détermination directe de leur valeur informationnelle.

Aspects technologiques et sociologiques de la communication

On peut se demander ce que signifie la communication: en d'autres mots, quelle est la différence entre la communication et l'information? Escarpit³³ affirme que la communication est un acte et que l'information est son produit. Cela veut dire que l'information est l'effet (un produit) d'un objet du troisième monde de Popper sur un objet du deuxième monde des états mentaux.

La technologie de la communication comprend tous les moyens de transmettre les messages, c'est-à-dire les livres, les journaux, les revues, les disques, les bandes magnétiques, les cassettes, les vidéocassettes, la radio, la télévision, le téléphone, la télégraphie, etc... Escarpit définit un document comme un objet informationnel visible ou touchable et doué d'une double indépendance par rapport au temps, à cause de sa synchronie et de sa stabilité. Les documents sont les objets physiques, ils ne sont pas une séquence d'événements. Escarpit fait une distinction entre les documents et les semi-documents. Les documents restent disponibles pour une lecture, libres de toute contrainte événementielle ou chronologique. Par contre, la perception des semi-documents de la technologie moderne comme par exemple les disques, les bandes, les films, les cassettes, etc. est liée à la disponibilité d'une reproduction du message original par des appareils.

Les moyens d'accès à l'information ont un aspect culturel aussi bien que technologique. Il y a bien sûr un autre troisième monde — le tiers-monde — et dans ce monde on perçoit souvent la plupart des moyens de la communication comme des armes de l'impérialisme culturel, des entraves à la communication libre et directe entre les individus, entre l'individu et l'État, et entre les États. Par exemple, à peu près 90% des nouvelles mondiales disséminées par les journaux, la radio, et la télévision en Asie, en Afrique, et en Amérique latine parviennent des centres de Paris, Londres, et New-York. Les sociétés multinationales comme IFT et IBM contrôlent l'industrie de

la communication et de la télématique. Par exemple, les États-Unis contrôlent environ 60-70% du marché pour l'exportation de l'information par ordinateur.³⁴

Le principe du flux libre de l'information entre les pays s'oppose souvent à celui de la défense de l'identité culturelle des pays moins développés. Ce conflit est parallèle à celui qui oppose le droit à l'information d'un individu et le droit à la confidentialité des autres individus. Les opinions divergent à ce sujet. Selon Du Long:

«Si ces moyens «d'information» ne sont pas adaptés à la sensibilité culturelle de la population cible, si l'arrière-plan culturel n'évolue pas dans le sens d'une utilisation accrue de ces moyens, c'est un coup d'épée dans l'eau...»³⁵

Un autre point de vue est celui de Choul:

«Si, comme l'affirme Ellul, la propagande présuppose l'information et une participation à la culture pour assurer son efficacité, il devrait être possible de les distinguer, mais dans nos sociétés surinformées une telle entreprise relève de l'utopie.»³⁶

Il est clair qu'à l'avenir les scientifiques de l'information ne devront pas se concentrer exclusivement sur des problèmes mathématiques et étroitement théoriques: il leur faudra faire la part des facteurs sociologiques, politiques et culturels. C'est à cette condition que la discipline conservera sa place au sein des sciences et augmentera l'importance de son apport à la pensée humaine.

RÉFÉRENCES

1. Suzanne Bertran-Gastaldy et Daniel Reicher, «La science de l'information à l'École de bibliothéconomie», *Revue canadienne des sciences de l'information*, vol. 5 (mai 1980), pp. 171-182.
2. B. Long, «Linguistique et indexation» *Documentaliste*, vol. 17, no. 2 (mars-avril 1980), pp. 99-106.
3. Karl Popper, *Objective Knowledge: An Evolutionary Approach*, (Oxford, Clarendon, 1973), 380 p.
4. B.C. Brookes, «The Foundations of Information Science. Part 1. Philosophical Aspects» *Journal of Information Science*, vol. 2, no. 3-4 (October 1980), pp. 125-133.
5. Gerard Salton and Michael J. McGill, *Introduction to Modern Information Retrieval*, (New York, McGraw-Hill, 1983), 448 p.
6. International Conference on Information Storage and Retrieval, *Proceedings, 1977-1982*, (Organisée par l'Association for Computing Machinery, SIGIR et les sociétés locales.)
7. C.J. Van Rijkbergen, *Information Retrieval*, 2nd ed., (London, Butterworths, 1979), 208 p.
8. Gilles Deschatelet, «Enquête sur les services de téléférence au Canada» *Documentaliste*,

- vol. 18, no. 6 (novembre-décembre 1981), pp. 207-213.
9. Charles Fillmore, «The Case for Cases» in E. Bach et R. Harms, ed. *Universals in Linguistic Theory* (New York: Holt, Rinehart, Winton, 1968), pp. 1-89.
 10. D. Lewis, «An Investigation of Case Grammar or Functional Relations in Language Expression as a Discriminating Factor in 'Aboutness' Recognition and Relevance Decision-making in Information Retrieval» Ph.D. Thesis proposal, University of Western Ontario, School of Library and Information science.
 11. J. Farradane, «The Nature of Informations» *Journal of Information Science*, vol. 1, no. 1 (April 1979), pp. 17-17.
 12. R.C. Cross, J.C. Gardin et F. Levy, *L'Automatisation des recherches documentaires* (Paris: Gauthier-Villars, 1964), 260 p.
 13. Avron Barr et E.A. Feigenbaum, *The Handbook of Artificial Intelligence* (Los Angeles: Kaufman, 1981), vol. 1, ch. 3-4.
 14. Y.A. Wilkes, «Preference Semantics» in E.L. Keenan, ed. *Formal Semantics of Natural Language* (Cambridge: Cambridge University Press, 1975), 475 p., pp. 329-348.
 15. R.G. Crawford, «The Relational Model in Information Retrieval» *Journal of the American Society for Information Science*, 1981, vol. 32, no. 1 (January 1981), pp. 51-64.
 16. Françoise Forest, «Une formalisation du comportement des différents interlocuteurs» *Documentaliste*, vol. 19, no. 1 (janvier-février 1982), pp. 16-19.
 17. Jean Claude Choul, «Le traitement de l'information dans un modèle de la lecture» *Revue canadienne des sciences de l'information*, vol. 7 (juin 1982), pp. 57-68.
 18. G.K. Zipf, *Human Behavior and the Principle of Least Effort* (Cambridge: Addison-Wesley, 1949), 373 p.
 19. Benoît Mandelbrot, «An Informational Theory of the Statistical Structure of Language» Dans Willis Jackson, ed., *Communication Theory* (London: Butterworths, 1953), 532 p., pp. 486-502.
 20. J.J. Hubert, «Linguistic Indicators» *Social Indicators Research*, vol. 3, no. 2 (June 1980), pp. 223-255.
 21. Jean Tague et M.J. Nelson, «Simulation of Information Retrieval Systems using Hyperterms» International Conference on Research and Development in Information Retrieval, West Berlin, 1982, *Proceedings* (sous presse).
 22. A.J. Lotka, «The Frequency Distribution of Scientific Productivity» *Journal of the Washington Academy of Sciences*, vol. 16 (1926), 323 p.
 23. S.C. Bradford, «Sources of Information of Special Subjects» *Engineering*, vol. 137 (January 1934), pp. 85-86.
 24. Derek de Solla Price, «A General Theory of Bibliometric and Other Cumulative Advantage Processes» *Journal of the American Society of Information Science*, vol. 27, no. 5-6 (September-October 1976), pp. 292-306.
 25. Eugene Garfield, *Citation Indexing - its Theory and Applications in Science, Technology, and Humanities* (New York: Wiley, 1979), 274 p.
 26. H.G. Small et D. Crane, «Specialties and Disciplines in Science and Social Science: an Examination of their Structure Using Citation Indexes» *Scientometrics*, vol. 1, no. 5-6 (August 1979), pp. 445-461.
 27. M.B. Lane et A. Sandison, «Obsolescence and Change in the Use of Literature with Times» *Journal of Documentation*, vol. 30, no. 3 (September 1974), pp. 283-350.
 28. B.C. Griffith et al., «The Aging of Scientific Literature: a Citation Analysis» *Journal of Documentation*, vol. 35, no. 3 (September 1979), pp. 179-196.
 29. C.E. Shannon et W. Weaver, *Théorie mathématique de la communication* (Paris: C.E.P.I., 1976), 388 p. (Traduit de l'anglais).
 30. Y. Bar-Hillel et R. Carnap, «Semantic Informations» *British Journal of Philosophic Science*, vol. 4 (August 1953), pp. 147-157.
 31. Y.A. Shreider, «On the Semantic Characteristics of Information» *Information Storage and Retrieval*, vol. 2, no. 4 (August 1965), pp. 221-233.
 32. Elisha Hoffman, «Defining Information: an Analysis of the Information Content of Documents» *Information Processing and Management*, vol. 16, no. 6 (1980), pp. 291-304; vol. 18, no. 3 (1982), pp. 133-139.
 33. R. Escarpit, *Théorie générale de l'information et de la communication* (Paris: Librairie Hachette, 1976), 218 p.
 34. Anthony Smith, *The Geopolitics of Information: How Western Culture Dominates the World* (Oxford: Oxford University Press, 1980), 192 p.
 35. Aubert DuLong, «Pour des études systématiques des politiques de développement des moyens d'information et de documentation» *Documentaliste*, vol. 19, no. 2 (mars-avril 1982), pp. 50-53.
 36. Jean-Claude Choul, «Obstacles à l'information: du support au contenu» *Comptes Rendus de l'Association canadienne des sciences de l'information* (vol. 9, 1981), pp. 120-131.



**CANEBSCO
SUBSCRIPTION
SERVICES, LTD.**



Six Thorncliffe Park Drive
Toronto, Canada M4H 1H3
(416) 421-9000

de **A** à **Z**

nous comblons tous vos
besoins d'abonnements

**LES SERVICES
D'ABONNEMENT
CANEBSCO LTÉE**

Notre représentant au Québec
M. Adrien B. Dumoulin
18 rue Thibault
Hull, Qué.
J9A 1H3 Tel: (819) 771-4237

Bibliographie rétrospective
Biblioservice multilingue
Canadiana
Catalogage avant publication
Catalogage lisible par machine
Centre de documentation sur les bibliothèques
Coordination des bibliothèques fédérales
Dépôt légal
ISBN
ISSN
Journaux
Listes collectives
Livres rares
Manuscrits littéraires
Musique
Prêt entre bibliothèques
Publications officielles
Référence et consultation
Réseaux
Services bibliographiques nationaux
Services pour les handicapés
Thèses canadiennes sur microfiches

**La Bibliothèque nationale du Canada
au service des bibliothèques canadiennes
d'un océan à l'autre.**



Bibliothèque nationale
du Canada

National Library
of Canada

Canada

La recherche en science de l'information dans les différents départements de l'Université de Montréal*

par Suzanne Bertrand-Gastaldy**

Une étude sur la recherche en science de l'information à l'Université de Montréal se heurte à divers obstacles. Ceux-ci sont d'abord d'ordre institutionnel: il n'existe pas de relevé systématique des activités de recherche menées dans l'organisme; aucune unité administrative ni aucun centre de recherche ne portent l'étiquette «Science de l'information». Ces obstacles constituent en fait les symptômes d'un mal plus profond de nature épistémologique: il est difficile de cerner le domaine de la science de l'information, de même que la notion d'information. Une première tentative pour repérer et regrouper les projets de recherche en science de l'information a cependant été tentée, en adoptant l'acception la plus large possible du mot information examiné sous une multitude de facettes. Quelques constats sont alors possibles. Des recommandations pour développer un véritable programme de recherche en science de l'information sont énoncées.

A study on research in information science at the University of Montréal encountered several obstacles. These are firstly of an institutional nature. There is no systematic listing of the research activities within the University. No administrative division or research center is identified by the name «information science». In fact, these obstacles are indicative of a more profound epistemological problem due to the difficulty of identifying the area of information science as well as the concept of information itself. An initial attempt to locate and regroup research projects in information science was undertaken. The largest possible definition of the word «information» was retained and it was examined under a number of aspects. The article presents some remarks as well as recommendations on developing a real research program in information science.

1. INTRODUCTION

Voulant vérifier si la science de l'information est, dans les faits, aussi interdisciplinaire, voire disparate qu'on a coutume de l'affirmer, nous nous sommes penchée sur la recherche en science de l'information dans un contexte précis mais assez large: l'Université de Montréal.

Comment avons-nous procédé pour mener notre étude? Essentiellement en consultant les sources documentaires mises fort aimablement à notre disposition par les services impliqués de l'Université (nous avons laissé de côté les écoles affiliées: Hautes Études Commerciales et Polytechni-

que). Après quelques remarques sur les limites imposées à une étude exhaustive, nous précisons la façon dont nous avons décidé d'aborder la science de l'information et particulièrement l'information. Nous illustrons ensuite nos constatations par des exemples de projets de recherches entrepris par différents départements. Enfin nous proposerons quelques actions visant à structurer la recherche dans ce domaine à l'Université.

2. LES OBSTACLES À UNE RECHERCHE SUR LA RECHERCHE EN SCIENCE DE L'INFORMATION À L'UNIVERSITÉ DE MONTRÉAL.

Les obstacles à une recherche sur la science de l'information à l'Université

de Montréal sont à la fois d'ordre institutionnel et d'ordre épistémologique.

2.1 Les obstacles d'ordre institutionnel

2.1.1 Les sources documentaires sur la recherche

Notre démarche nous a permis de faire quelques constatations concernant les sources documentaires sur la recherche à l'Université de Montréal.

Ces sources:

- sont diverses (deux unités s'en occupent principalement: Service de la recherche, Direction des communications, mais certaines facultés et certains départements font leur propre relevé);
- ne couvrent pas les mêmes type de recherche (recherche subventionnée uniquement ou toute ac-

* Texte reçu de la conférence prononcée le 8 avril 1983.

** Suzanne Bertrand-Gastaldy est professeur agrégé à l'École de bibliothéconomie de l'Université de Montréal.

tivité conduisant à une publication);

- ne sont pas conçues en fonction des mêmes objectifs;
- ne contiennent pas les mêmes éléments d'information;
- de façon générale, ne fournissent que très peu de renseignements sur le contenu.

En fait, il faut bien l'avouer, il n'existe pas de contrôle bibliographique de la recherche à l'échelle de toute l'Université. Les instruments existants sont orientés surtout vers les besoins administratifs (ils permettent par exemple d'obtenir des statistiques sur les subventions obtenues par département) et, dans une certaine mesure, vers les relations publiques (par le truchement du bulletin *Recherches à l'Université de Montréal* destiné surtout aux journalistes). Le Service de la recherche avait entrepris la compilation d'un inventaire des publications de recherche. Celle-ci a été interrompue pour des raisons économiques et reprise partiellement par d'autres unités.

L'Université compte apparemment sur l'efficacité des échanges informels pour la circulation de l'information sur la recherche.

Mais, de ce fait, aucune étude exhaustive sur la recherche sur le plan synchronique ou diachronique n'est possible.

2.1.2 L'absence d'étiquette «science de l'information» dans la structure de l'Université de Montréal

Aucun département, aucune faculté, aucun centre de recherche à l'Université de Montréal ne portent actuellement l'étiquette «science de l'information».

Si des activités se déroulaient dans ce domaine, ce serait sinon à l'insu de tous, du moins de façon inconsciente. Quelques personnes – ou plusieurs (ce sera à notre enquête d'apporter des éléments de réponse) – feraient peut-être de la science de l'information sans le savoir ou sans que la communauté universitaire le sache!

Nous sommes quelques-uns à connaître le projet de l'École de bibliothéconomie qui est en voie de changer son nom pour celui d'École de bibliothéconomie et de science de l'information, à l'instar de la plupart de ses homologues en Amérique du Nord.

Mais à constater l'étonnement et les réticences de certains, force est d'admettre qu'on ne lui accorde pas d'emblée l'exclusivité en ce domaine.

Ce malaise n'est pas imputable à l'ignorance des responsables universitaires, mais constitue plutôt le symptôme d'un mal plus profond et généralisé.

En effet, des experts de six pays réunis en 1979, à une table ronde sur les programmes nationaux de recherche en science de l'information, à Georgie Tech's School of Information and Computer Science exprimaient l'impérieuse nécessité de structurer la recherche et révélaient du même coup que la science de l'information n'est pas encore reconnue comme une discipline à part entière faisant l'objet d'un enseignement spécifique rattaché à une seule unité.¹ Beaucoup d'universités sont encore structurées selon les disciplines traditionnelles et continuent de fonctionner comme si nous n'étions pas encore entrés dans l'ère de l'information – qui pourtant bouleverse nos modes de vie. Ainsi la «non-visibilité» de la science de l'information dans les universités résulterait non pas tant du conservatisme de nos vénérables institutions que de la difficulté éprouvée – souvent par les chercheurs eux-mêmes – à circonscrire le domaine de même que la notion d'information.

2.2 Les obstacles d'ordre épistémologique

2.2.1 Absence de délimitation claire du domaine

C'est un lieu commun que de s'interroger sur le contenu de la science de l'information, sur son caractère scientifique et sur son interdisciplinarité.

Elle est considérée tantôt comme une discipline en voie de formation, tantôt comme un conglomérat de disciplines, sans identité propre, à tel point qu'on n'ose pas toujours la nommer au singulier:

The first problem that we met with this subject is how to define Information Science. It is quite significant to note that in French we speak of «Science de l'Information», which clearly makes out that these are included and these together constitute a kind of discipline.

In fact the information sciences appear to be at the meeting point of other sciences like computer science, linguistics, electronics, telecommunication, mathematics, psychology, sociology, economy, etc., and it took quite a long time for Information Sciences to acquire their identity, their autonomy.²

Pourtant, après quelques précautions oratoires, de bon aloi, personne n'hésite à proposer sa propre définition. L'une d'entre elles a cours depuis le début des années 1960 et est fréquemment citée:

The science that investigates the properties and behavior of information, and the means of processing information for optimum accessibility and usability. The processes include the organization, dissemination, collection, storage, retrieval, interpretation and use of information.³

Bien que relativement complète, cette définition a le défaut souligné précédemment par Jean Tague:⁴ elle ne précise pas la signification du mot information.

2.2.2 Polysémie du mot information

Or, peut-on prétendre cerner la science de l'information tant qu'on n'a pas cerné l'information?

À la limite tout est information et tout le monde peut prétendre s'en occuper:

The field of possible study becomes so wide as to be all embracing. Over 2000 years ago Heraclitus said that: "All is flux" and now it may seem that I am saying that: "All is information". (...)

There seems to be a great deal of overlapping both in basic theory and in the use of language but on the other hand there are a great number of different aspects of the same problem being discussed. One notable difference is in the interpretation of words; the most disturbing is the use of the word "information" which seems to be the common denominator although it means so many different things to different people.⁵

C'est probablement la polysémie du mot information qui explique le caractère souvent hétéroclite:

- des listes de disciplines accompagnant les tentatives de définition
- des sujets traités dans les publications (périodiques, comptes rendus, manuels, etc.)
- des programmes d'études – des politiques nationales
- des centres d'intérêt des associations, des institutions d'enseignement et de tous les organismes qui s'y consacrent.

Il faut peut-être voir aussi dans cette polysémie l'origine du savoir démesuré qu'on exige parfois du spécialiste en science de l'information qui devrait maîtriser une vingtaine de disciplines à la fois, ce qui ferait de lui l'humaniste des temps modernes.⁶

2.2.3 La solution: ouvrir les frontières

Bien que l'on s'accorde à reconnaître que le champ d'application de la science de l'information, ses principes théoriques, ses concepts fondamentaux et ses méthodologies sont mouvants, on n'en prône pas moins d'aller de l'avant.

Les mots qui reviennent le plus souvent: «interdisciplinarité», «coopération», «interaction» évoquent assez la volonté de trouver dans l'action con-

certée la solution, quitte à reporter les problèmes de définition dans le futur.

Les suggestions d'ouvrir les frontières plutôt que d'en imposer d'artificielles, de transformer la science de l'information en forum de recherche critique plutôt que d'en faire une chasse gardée, d'accepter la polysémie des mots plutôt que d'en réduire l'acception, sont de plus en plus écoutées et mises en application.

... whether we call ourselves library scientists, computer scientists, or communication scientists, we are all information scientists concerned with the communication of information. Our interests differ only in details such as transmission medium, message content, types of sources and receivers, and persuasive intent. But we rarely communicate with each other. I am merely suggesting first, that we should be prepared to consider the wider uses of the term *information* before we commit ourselves to research on any foundational theory or before we stipulate any restricted use of the term for our particular purposes, and, secondly, that we must respect other uses of the term *information* if we hope to open information science to free and critical scientific discussion.⁵

Et l'on pourrait suivre les conseils d'un auteur confronté au même phénomène, mais à propos de la théorie de la communication:

La première étape de l'élaboration d'une théorie de la communication consiste à recenser minutieusement les différents domaines dans lesquels ce concept a été appliqué: le langage verbal ou non verbal, les médias et la culture, les télécommunications et l'informatique.

Ne pas craindre le disparate et l'hétérogénéité. Voir ce qu'apporte l'expérience de chacun de ces domaines, en quoi elle enrichit le concept de communication, quels sont les aspects de ce concept qu'elle met en valeur, et ceux qu'elle laisse dans l'ombre. Analyser dans chacun de ces domaines les formalisations théoriques auxquelles il a donné naissance. Examiner en quoi chaque formalisation tire parti d'une expérience particulière, et dans quelle mesure elle est infléchi par des contraintes spécifiques.

Rechercher finalement s'il existe un concept central qui rende compte de l'ensemble de ces pratiques, se demander en somme si le temps est venu d'une théorie générale de la communication.⁶

Remplaçons le mot communication par le mot information et nous avons là un programme d'action tout tracé.

Il nous faudra donc inventorier les différentes facettes du concept d'information avant de repérer les projets de recherche qui traitent de l'une ou de plusieurs d'entre elles.

2.2.4 Les diverses facettes de l'information

L'information est omniprésente, nous l'avons dit.

Selon le niveau auquel on se situe, il s'agit d'information génétique (l'émetteur et le récepteur étant la cellule) ou de savoir, c'est-à-dire de données assimilées, enregistrées et retransmises par les êtres humains.

Cette information possède certaines propriétés: probabilité, pertinence, comptabilité, exhaustivité, redondance, complexité, désuétude, etc.

Elle est l'objet de certaines activités: génération, acquisition, organisation, exploitation, transmission, diffusion.

Elle obéit à certains phénomènes: compréhension, rétention, oubli.

Elle est mémorisée de façon individuelle (par la cellule, par le cerveau) ou institutionnelle (centres de documentation, photothèques, bibliothèques, etc.).

Elle peut être consommée immédiatement ou réutilisée en différé.

Elle est consignée sur un support matériel: tablette d'argile, papier, tissu, support biologique, support magnétique, pellicule, etc.)

Son enregistrement sur ce support fait appel à diverses techniques: imprimerie, physique (optique), informatique, ingénierie électronique, radiodiffusion, télécommunication.

Elle est représentée par des codes aux multiples combinaisons: langage oral ou écrit, visuel ou sonore, binaire ou autre, selon différents canaux de transmissions: ondes hertziennes, impulsions électriques, donnant lieu à une multitude de documents manuscrits, imprimés, visuels, audios, audiovisuels, ordnolingués.

Toutes ces caractéristiques de l'information peuvent être envisagées sous des angles divers: temporels (délais), historiques, économiques, sociologiques, légaux, sémiologiques, psychologiques, technologiques, quantitatifs.

Les êtres humains qui interviennent dans le processus de la communication de l'information peuvent être catégorisés:

- selon leur âge (enfants, adolescents, adultes, personnes âgées),
- selon leur activité professionnelle (écoliers, étudiants, enseignants, chercheurs, administrateurs, techniciens)
- selon leur rôle social (consommateurs, citoyens)
- selon leur statut juridique (prisonniers)
- selon leur état de santé (convalescents, malades, handicapés physiques)
- selon la place qu'il occupent dans le schéma de la communication

(émetteurs, récepteurs, intermédiaires)

Les buts de l'émission ou de l'utilisation de l'information peuvent être: la publicité ou la consommation, les relations publiques, la propagande, l'instruction ou l'éducation, la prise de décision, le divertissement, la culture.

Son objet peut être le domaine scientifique, technique, artistique, culturel.

Si nous considérons l'information sous toutes ces facettes, alors nous ne nous étonnerons plus qu'un si grand nombre de disciplines se réclament de la science de l'information et que tant de spécialistes soient susceptibles d'intervenir à un titre ou à un autre dans son développement.

L'information (ainsi conçue) reste information quels que soient son objet, son but et sa forme matérielle. C'est-à-dire qu'il n'y a pas lieu de faire une distinction entre l'information au sens habituel et journalistique du mot, le commentaire, la propagande, la publicité, la matière des relations publiques, les distractions, l'instruction, l'éducation et même ce que chacun peut recueillir par ses contacts personnels et son expérience.⁷

Une démarche aussi éclectique a l'avantage d'éliminer d'emblée plusieurs contresens:

- information = documents
- information = documentation scientifique et technique
- science de l'information
- = informatique
- = technologie
- = méthodes quantitatives

Entre les spécialistes en information tout court et les spécialistes en information documentaire, il n'y aurait qu'une différence de degré. Ces derniers seraient concernés par la fourniture, à tout usager, quelle que soit l'utilisation qu'il veuille en faire, de toute information, pourvu que celle-ci ait été enregistrée, traitée, analysée et repérée, quel que soit le support et quel que soit le médium.

Mais cette information enregistrée est le produit d'une pensée humaine qui modifie le savoir d'une autre personne. Donc l'information documentaire ne peut évoluer sans que soient étudiés et compris les phénomènes plus globaux de la cognition.

Et l'on a probablement tout à gagner en généralisant l'étude de l'information jusqu'à inclure l'examen du fonctionnement particulier des systèmes biologiques.

C'est donc en abolissant les frontières entre les différentes disciplines, entre les différentes facettes de l'information que l'on pourra vraiment aboutir à une amélioration du trans-

fert des connaissances et à une véritable théorie de l'information.

On doit faire tomber la distinction entre l'information véhiculée par les documents, par les mass média, par les êtres humains, par la cellule même.

3. LA SCIENCE DE L'INFORMATION À L'UNIVERSITÉ DE MONTRÉAL

L'information ainsi décloisonnée fait l'objet de plusieurs dizaines de recherches à l'Université de Montréal.

Compte tenu de l'importance du nombre et des conditions — mentionnées au début — dans lesquelles s'est effectuée cette enquête, il est impossible de rendre compte de chacun des projets sans risquer de ne pas faire justice à tous.

Cependant quelques lignes de force se dégagent et quelques constatations s'imposent que nous illustrerons par des exemples

3.1 Il y a plus de points communs que de différences entre la plupart des études menées dans les différents départements

— Prenons le cas de:

J.K. Szpakowska. *Besoins en information et en services documentaires chez les jeunes de 14-19 ans*. Sondage réalisé à la Bibliothèque municipale de Ville Saint-Laurent, novembre 1977 (École de bibliothéconomie).

et de: D. Brady et B. Schiele. «Développement et évaluation des modèles d'émissions télévisuelles d'informations sociopolitiques et de vulgarisation scientifique destinés aux adolescents», (département de technologie éducative de la faculté de science de l'Éducation).

Quel que soit le médium, il s'agit de se préoccuper des besoins en information des adolescents, dans divers milieux: à la bibliothèque, chez eux, en classe, etc.

— ou bien:

A. Méar. «Élaboration d'une méthodologie pour l'analyse de l'utilisation que font les canadiens de la télévision» (Département de communication)

et: J. de Bonville. *Les habitudes documentaires des journalistes des quotidiens francophones du Québec, mai 1977*, 215 p. (Mémoire supervisé par Marcel Lajeunesse, à l'École de bibliothéconomie).

Adolescents, enfants, étudiants, grand public, personnes handicapées, consommateurs, tous les utilisateurs

d'information sont étudiés par les trois départements cités précédemment et par d'autres comme le groupe de recherche en consommation (C. Masse, «L'information et l'exploitation des consommateurs») avec des nuances (par rapport au médium, au type d'information, au genre d'utilisation).

3.2 Certains sujets étudiés dans plusieurs départements constituent cependant des points forts de quelques-uns d'entre eux

Les archives font l'objet des préoccupations du département d'histoire (J. Rouillard, «La récupération des archives syndicales au Québec»), du département de criminologie (J. Trépanier, «Les archives des prisons québécoises»), du service des archives (C. Couture et Y. Gendreau, «Bibliographie des archives»), alors que le livre pour enfants, les dessins animés, les documents audiovisuels sont analysés et évalués en fonction de leur impact pédagogique au département de technologie institutionnelle. Le département des communications s'occupe, mais il n'est pas le seul, surtout des médias télécommunicateurs. La Faculté d'aménagement est la seule unité à avoir entrepris des recherches sur «la micrographie informatique dans la presse scientifique». (C.H. Davidson et M. Gagné).

Aucun département n'a l'apanage de la création de bases de données numériques ou bibliographiques, d'index, de répertoires de catalogues, de registres, d'inventaires analytiques, de dictionnaires, établis à la main ou par ordinateur.

Sont impliqués aussi bien les Facultés d'aménagement, des sciences de l'Éducation, de droit que les départements de bibliothéconomie, de démographie, d'histoire, d'études médiévales, de géographie et le groupe de recherche en développement économique, à la Faculté des Arts et des sciences.

Il s'agit cependant d'expériences sporadiques et pratiques. La concentration des recherches qui font progresser les méthodes de création et de gestion de bases de données se trouve bien sûr au département d'Informatique et de recherche opérationnelle, tandis que les méthodes de conception et surtout d'évaluation d'outils ou de systèmes bibliographiques de repérage manuel ou automatisé sont localisées à l'École de bibliothéconomie.

3.3 L'ensemble des activités de recherche donne plus l'impression d'une juxtaposition d'actions ponctuelles que d'une intégration d'actions concertées

Les aspects économiques, politiques, culturels et sociaux de l'information sont tous étudiés, mais de façon cloisonnée, à l'intérieur du champ de spécialisation de chaque département.

Le centre de recherche en droit public s'intéresse au droit du public à l'information, à la propriété intellectuelle des jugements, au droit d'auteur en éducation, à la protection des brevets d'invention et des transferts de technologie.

Le département des sciences économiques a entrepris une comparaison interrégionale de la diffusion des innovations au Canada (I. Martin) tandis que le département des sciences politiques a pour projet: «Technologie et politique au Québec: évaluation des transferts technologiques et perspectives de développement économique» (G. Ducatenzeller et G. Boismenu).

La diffusion du savoir scientifique, dans la perspective politique et historique, sa vulgarisation relèvent, bien sûr, du département d'histoire et de socio-politique des sciences. Un de ses professeurs, Y. Rabkin, effectue la mesure de cette diffusion dans les revues scientifiques.

Pour sa part, le département des communications a procédé à une «Évaluation critique des politiques du gouvernement québécois dans le domaine des communications» et, au département de technologie institutionnelle de la Faculté des sciences de l'éducation, on relève le projet suivant: «A study of the economic, policy and institutional issues influencing the use of satellite based telecommunications facilities for public service» (H.D. Stolovitch, P. Ferusse, M. Kassner et D.P. Brassar).

Dans un même ordre d'idée, Marcel Lajeunesse de bibliothéconomie, met en lumière l'histoire de la diffusion des idées, du XIXe au XXe siècles et le rôle social et culturel qu'ont joué les associations littéraires, les conférences publiques, les cabinets de lecture, les bibliothèques.

C'est le groupe GAMMA, avec K. Valaskakis et P. Simoell, qui a peut-être la vision la plus large: «Implications socio-économiques de la société informatisée», «Analyse de certaines tendances dans le secteur informationnel et des implications pour la société

canadienne»; «Les impacts sectoriels d'une société informatisée» (K. Valaskakis).

3.4 L'interdisciplinarité, nécessaire pour étudier certains domaines en émergence, est plus le fait d'initiatives individuelles que d'équipes de recherches dûment constituées en vue d'une mission

Certains ont franchi quelques-unes des barrières dressées entre les disciplines. D. Sankoff, du Centre de recherche en mathématiques appliquées, s'intéresse aux problèmes de linguistique mathématique et de biologie mathématique, aux applications de la classification biologique et linguistique.

P. Rolland-Thomas, de l'École de bibliothéconomie, propose, pour établir la validité des classifications documentaires, de les situer à la fois dans l'ordre de la culture (développement des classifications dans la langue, l'organisation sociale, les institutions, les cultes, l'épistémologie des connaissances), de la nature biologique (la façon dont la cellule reçoit et sélectionne l'information pour sa survie) et de la psychologie (la façon dont l'être humain développe ses aptitudes à la classification du monde concret et abstrait).

— Autre exemple, lié cette fois au problème de la représentation de l'informatique, au langage.

Informatique, statistique, linguistique quantitative, sémantique, syntaxe, etc... se retrouvent dans bien des projets, avec une forte concentration autour de S. Lusignan (Institut d'études médiévales) et P. Bratley (Informatique et recherche opérationnelle) et du groupe de recherche en traduction automatique de l'Université de Montréal (TAUM), maintenant dissout.

Il n'y a probablement pas une grande différence entre:

d'une part le projet de G. Allard (Institut d'études médiévales) sur le «De Divisione Naturae» de Jean Scot Erigène qui étudie le texte, son vocabulaire, la syntaxe, les interrelations d'idées, la stratégie rhétorique, pour mieux comprendre les processus argumentatifs en philosophie

et d'autre part le traitement des analyses du *Bulletin Signalétique 101* par S. Bertrand-Gastaldy et Y. Henry-Rousseau (Bibliothéconomie) qui étudient la répartition des mots, des vocables, des radicaux, des expressions composées. Leur intention est de sélectionner les éléments les plus pertinents

du texte et de les regrouper en fonction de leurs relations syntagmatiques et paradigmatiques pour offrir à l'utilisateur de la base de données un outil de repérage de l'information en langage naturel.

Seule la finalité des deux études change.

Les chercheurs observent presque les mêmes phénomènes et utilisent presque la même méthode. D'ailleurs les deux groupes se servent du langage de programmation DEREDEC mis au point par Pierre Plante, analyste à l'UQAM.

— L'étude de L. Thibodeau-Brunet (École de bibliothéconomie), pour la mise au point d'un programme de formation documentaire des étudiants avec le service de pédagogie de l'UQAM, est menée en parallèle avec une équipe de psychologues travaillant sur la métacognition (observation des processus qui interviennent dans l'acte de connaître et qui sont liés, entre autres, aux habitudes de lecture).

Or, deux chercheurs de la Faculté des sciences de l'éducation étudient eux aussi «Les processus cognitifs de l'étudiant universitaire» (S. Racine et R. Larose).

Lorsque L. Thibodeau-Brunet se propose de mener une enquête de type psychosocial afin d'évaluer la satisfaction des besoins d'information des handicapés face au vidéotex (en terme de type d'informations sélectionnées et d'organisation d'information), elle rejoint le propos de Mills et Giroux, du département des communications qui se demandent comment structurer l'information pour mieux répondre au processus cognitif humain.

Et parce que le vidéotex fait partie de ces nouveaux média qui permettent à l'utilisateur de jouer un rôle actif (alors que la télévision, la radio et le cinéma le laissent passif), un peu comme il peut se comporter avec les média traditionnels comme le livre, la finalité des chercheurs de différentes disciplines devient la même: faciliter le repérage de l'information, par tous les moyens possibles, psychologiques aussi bien que techniques.

3.5 Un exemple de domaine interdisciplinaire en émergence: la psychologie ergonomique

On peut juger de l'interdisciplinarité des recherches sur le vidéotex par les quelques projets suivants:

— Au département des communications:

M. Mills: "Study of the Human

Response to Pictorial Representations on Telidon"

"Can Outline Pictures substitute for Photographs on Telidon? Correlating Attitudes with Performance".

"Behavioral Research on Telidon (Proposal for a Working Meeting)"

L. Giroux: «Processus cognitifs et recherche d'information sur vidéotex».

— Au département d'informatique et de recherche opérationnelle: G. Bochmann et J. Gecsei: «Coopération de bases de données réparties pour les applications Vidéotex (Intworking and Distribution of Data Bases for Videotex Applications)"

"Data Base Management Systems for Telidon and Videotex Information Systems".

"Enhancement of Videotex by use of data base machine and special purpose hardware".

Inclure les principes et méthodes des psychologues est indispensable pour la conception de nouveaux modes de repérage, de nouvelles stratégies de recherche vraiment adaptées au processus cognitif des utilisateurs, de même que pour l'évaluation des systèmes et outils d'information.

Cependant la psychologie ergonomique en est encore à ses débuts.

Les systèmes documentaires actuels n'ont pas été et ne sont toujours pas conçus sur la base de recherches ergonomiques, leurs concepteurs sont orientés par les problèmes, alors qu'ils devraient l'être par l'opérateur. C'est l'analyse de l'activité humaine réelle et quotidienne face au problème qui devrait être la base primordiale de la conception d'un système. Des méthodes existent, celles de la psychologie expérimentale, et un corps de connaissances qu'il faut faire progresser, mais qui pourraient déjà être mieux exploitées.¹³

Donc l'intégration des recherches psychologiques, bien sûr fort bien représentées au département de psychologie et aboutissant parfois à l'intelligence artificielle, permettrait de renforcer cet intérêt de plus en plus fort pour les utilisateurs et les agents du transfert de l'information.

Justement l'équipe Davidson, Crapo et Bertrand-Gastaldy, dans son projet de «Méthodes de compression de thésaurus en vue d'optimiser le repérage de l'information», s'intéresse à des questions du genre: quelle est la quantité optimum de descripteurs à présenter sur un écran? quelle est la meilleure représentation graphique des relations? Comment permettre à chaque

individu de se «promener» dans le langage documentaire en fonction de son degré de familiarisation avec le système?

En analyse documentaire, il est de toute première importance de comprendre comment se forment les concepts, comment ils sont mémorisés ou rejetés, interprétés, comment s'effectue la prise de décision, puisque le documentaliste ou le bibliothécaire, constitue une barrière supplémentaire entre la source d'information et son récepteur, étant lui-même un récepteur et transmetteur intermédiaire.

L'interface homme-information aboutit tout naturellement à l'interface homme-machine (psychologie, documentation, informatique, etc.)

L'École de bibliothéconomie a fait de l'utilisateur le centre de son programme d'études et l'on constate que les recherches ne visent plus tellement à mesurer quelle utilisation il fait des objets tangibles (documents, bibliothèques, ordinateurs), mais bien comment il interagit avec l'information et comment on pourrait lever les obstacles qui freinent l'utilisation de l'information.

L'accent est également mis sur cet autre obstacle à la diffusion de l'information que peut constituer l'agent de transfert de l'information.

Dans l'étude des caractéristiques socio-culturelles des bibliothécaires, R. Savard s'intéresse aux facteurs qui peuvent influencer la qualité et l'efficacité de la relation bibliothécaire-client, et particulièrement dans quelle mesure les attitudes professionnelles, comme l'orientation service, peuvent jouer un rôle dans le transfert de l'information.

A. Perrier, lui, poursuit une recherche sur «Le rôle des facteurs culturels (langue, scolarité, expérience) dans l'évaluation de la performance des bibliothécaires en rapport avec leur capacité à résoudre des questions de référence: comparaison entre bibliothécaires britanniques et québécois».

3.6 La place de l'École de bibliothéconomie dans la recherche en science de l'information à l'Université de Montréal

Les quelques projets mentionnés précédemment montrent bien que l'École de bibliothéconomie a suivi les conseils de Wright qui incitait les spé-

cialistes de l'information à décider s'ils administreraient des choses ou plutôt la pensée, l'information, les idées, le savoir.¹²

L'École en effet ne considère plus l'information comme une commodité seulement et ne la confond plus avec les documents qui la contiennent (archives, livres, banques de données). Elle ne s'intéresse plus uniquement aux solutions pratiques, aux aspects techniques, aux opérations (génération, dissémination, sélection, analyse et représentation, stockage, recherche), aux instruments: à la main d'œuvre (éducation, qualification), à la méthodologie (langages d'indexation, stratégies de recherche).¹³

C'est dans la mesure où elle réussira à mettre l'accent sur l'information en tant que processus, c'est-à-dire comme une modification de l'état mental d'un individu (par l'apport d'un contenu), ainsi que le mentionnait Jean Tague, qu'elle méritera vraiment de s'appeler École de bibliothéconomie et de science de l'information.

Elle devra privilégier l'étude de l'esprit, la façon dont celui-ci est affecté par l'information (par sa quantité, sa présentation, sa nouveauté, sa nature, le type de médium qui la véhicule, etc.), la façon dont l'homme — et la société — l'accumule, l'évalue, la traite, l'oublie, la rejette.

L'École peut-elle à elle seule poursuivre des objectifs aussi ambitieux? Certes pas.

D'abord elle n'a pas les ressources nécessaires.

On peut regretter l'absence de recherches sur certains sujets qui devraient l'intéresser au premier chef: le droit d'auteur, l'information iconographique, par exemple, et la timidité de certaines initiatives dans des secteurs neufs.

Avec quelques professeurs-chercheurs supplémentaires, elle pourrait probablement pallier plusieurs lacunes.

En collaboration avec des chercheurs d'autres départements ou facultés (informatique, psychologie, linguistique, communication, histoire et sociopolitique des sciences, sciences politiques, sciences économiques, Faculté de droit), elle pourrait offrir des champs d'application pratique à des théories en mal de validation, et contribuer à mieux comprendre les rapports entre l'homme et l'information, entre la société et l'information, elle qui a une longue expérience de l'infor-

mation enregistrée sur des supports documentaires.

4. POUR UNE PLANIFICATION DE LA RECHERCHE EN SCIENCE DE L'INFORMATION À L'UNIVERSITÉ DE MONTRÉAL

En fait, ce qu'il manque à l'Université, c'est une volonté commune des chercheurs de divers horizons et la conviction qu'ensemble ils pourraient réfléchir sur ce qui est peut-être le problème le plus fondamental de l'époque.

Mais bien sûr, il y a plusieurs étapes à franchir avant d'aboutir à une véritable concertation.

Il faudrait envisager rapidement des mesures concrètes et systématiques pour favoriser des échanges entre personnes isolées dans leur champ de spécialisation. Puisque les ressources humaines sont là — nous l'avons perçu de façon non exhaustive mais de façon suffisamment claire au cours de notre enquête —, il reste à les exploiter.

— D'abord, un inventaire minutieux des recherches menées sur l'information s'impose.

— Cet inventaire devrait conduire tout naturellement à recenser tous les chercheurs impliqués.

— Des rencontres informelles autour d'un problème spécifique devraient être organisées afin qu'on puisse entendre le point de vue de chaque discipline, afin aussi que les chercheurs apprennent à se connaître.

— Il ne serait alors plus utopique d'envisager l'élaboration d'une politique de la recherche en science de l'information à l'Université de Montréal.

— De là, on pourrait songer à la création d'un centre ou de groupes de recherche.

— Munis d'un programme, de priorités, de structures internes assurant le soutien à la recherche, il serait plus aisé pour les chercheurs d'obtenir des subventions propres à garantir un développement continu du domaine.

Nul doute qu'ainsi organisés, les chercheurs parviendront à asseoir la science de l'information sur des bases théoriques solides.

Si rien de tout cela n'existe actuellement, il faut nous hâter, avant que d'autres ne décident à notre place.

En effet c'est peut-être une chance qu'il n'existe ni au Québec ni au Canada de programme national de scien-

ce de l'information. Car, comme le faisait remarquer Slamecka,¹¹ parce que les chercheurs en science de l'information n'ont pas su, contrairement à leurs collègues des autres disciplines, définir leurs propres priorités, les gouvernements se sont substitués à eux et ont imposé des programmes centrés sur le développement de grands systèmes d'information, répondant à des objectifs très pragmatiques qui ne correspondent pas forcément aux préoccupations des chercheurs.

D'après cet auteur, seuls les États-Unis ont des priorités qui ne sont pas uniquement orientées vers des aspects pratiques et encouragent la recherche fondamentale.

Lorsqu'on voudra planifier la recherche à l'Université de Montréal, on pourra d'ailleurs se référer à la liste des projets subventionnés par la National Science Foundation et qui se répartissent en quatre grands thèmes: normes et mesures, structure de l'information, aspects comportementaux du transfert de l'information, économie de l'information.¹²

5. CONCLUSION

Voilà donc un survol bien rapide de la science de l'information à l'Université de Montréal. Il faut le redire, l'enquête qui nous a servi de base est superficielle. Souhaitons qu'on puisse un jour y consacrer les ressources nécessaires et bénéficier de l'apport de tous les chercheurs avec la coordination de l'Université.

En attendant, on peut d'ores et déjà conclure que:

— Le domaine de la science de l'information est trop vaste pour être délimité d'emblée.

— Beaucoup plus de chercheurs qu'on pourrait le penser à première vue se livrent à des recherches sur divers aspects de l'information.

— En l'absence de plan d'ensemble, ces recherches s'effectuent souvent en vase clos. Ce cloisonnement vertical est préjudiciable à la science de l'information qui, on l'a vu, est avant tout interdisciplinaire.

— La richesse des études déjà réalisées, en cours de réalisation ou en projet, serait décuplée si elles faisaient l'objet d'une véritable concertation. Les études potentielles sont nombreuses.

— L'École de bibliothéconomie est dans la bonne direction, et c'est peut-être sous son instigation que pourraient être rassemblés les éléments disparates du puzzle.

NOTES INFRAPAGINALES

1. Les textes de ces interventions sont réunis dans un numéro spécial de *Information Processing & Management*, vol. 16, (1980), pp. 177-257.
2. Serge Chambaud et Thérèse Delpech, «Research in Information Science in France», *Information Processing & Management*, vol. 16 (1980), p. 194.
3. Cette définition vient des Conférences du Georgia Institute of Technology qui ont eu lieu en octobre 1961 et avril 1962, d'après Jesse H. Shera et Donald B. Cleveland, «History and Foundations of Information Sciences», *Annual*

Review of Information Science and Technology, vol. 12 (1977), p. 264.

4. Jean Tague, «Information Science in Graduate Library Programs», *Canadian Library Journal*, vol. 36, no. 3 (June 1979), p. 39.

5. Fritz Machlup, «Sciences of Information. Looking over the Fences», *Communicating Information. Proceedings of the 4th ASIS Annual Meeting*, vol. 17 (1980), p. 6.

6. Louis Vagianos, «Information Science: a House Built on Sand», *Library Journal*, vol. 97, no. 2 (January 15, 1972), p. 153.

7. Thomas J. Kirwin III, «A Communicator's View of Information Science», *Bulletin of the American Society for Information Science*, vol. 3, no. 6 (August 1977), p. 24.

8. B.C. Brookes, «The Fundamental Problem of Information Science», *Informatics 2: Proceedings of a Conference held by the ASIS Co-ordinate Indexing Group on 25-27 March 1974 at New College, Oxford* (1974), p. 49.

9. Jacques Durand, *Les formes de la communication* (Paris, Dunod, 1981), p. 1.

10. P.M.G. Levy, *Revue des sciences sociales de la France de l'Est*, 1, 1972, cité dans J. Weber-Amouyal, *Information et pouvoirs* (Paris, Hatier, 1981), p. 8. (Coll. Profil Dossier, 534).

11. A. Bisseret, «Pour une psychologie ergonomique des systèmes documentaires», *Documentaliste*, vol. 20, no 1 (janvier-février 1983), résumé p. 6.

12. H. Curtis Wright, «The Wrong Way to Go», *Journal of the American Society for Information Science*, vol. 30, no. 2 (March 1979), p. 67.

13. Y. Le Coadic, «Quelques aspects de la science de l'information en 1977», *Documentaliste*, vol. 15, no 3 (juillet 1978).

14. Vladimir Slamecka, «National Agendas of Research in Information Science; an Overview», *Information Processing & Management*, vol. 16 (1980) pp. 252-253.

15. Howard L. Resnikoff, «Information Science and Technology at the National Science Foundations», *Information Processing & Management*, vol. 16, (1980), pp. 243-250.

UTLAS

CONNAISSEZ-VOUS UTLAS?

Nous vous connaissons déjà car voilà plus d'une décennie qu'**UTLAS** dispense au Québec des services automatisés, en français et en anglais, à des bibliothèques publiques, académiques, et spécialisées.

Permettez-nous donc de nous présenter ...

"À TITRE D'INFORMATISATION"

Services offerts:

- Système de soutien de catalogage CATSS
- Contrôle des autorités (noms et matières)
- Conversion rétrospective (RECON) des fichiers manuels
- Contrôle des acquisitions
- REFCATSS et module de prêt entre bibliothèques
- Boîte aux lettres électronique
- Contrôle des publications en série
- Possibilités d'interface avec d'autres systèmes

Produits:

- Microcatalogues et catalogues imprimés sur feuilles
- Fiches de catalogue
- Listes de nouveautés
- Épreuves de catalogage
- Index permutés du type KWIC
- Étiquettes, bons de commande, et rapports de gestion générés en bibliothèque
- Bandes magnétiques

Maintenant que vous nous connaissez un peu mieux, communiquez donc avec notre Service du marketing. Nous sommes situés au 80, rue Bloor Ouest, Toronto (Ontario) M5S 2V1; ou téléphonez-nous au 416/923-0890.

UTLAS ... AU SERVICE DES BIBLIOTHÈQUES DU MONDE ENTIER

Modelling Problems in Bibliometrics

By John J. Hubert*

Several recent examples from library stock management, citation analysis, linguistics and productivity modelling are used to illustrate problems which have been solved and problems which are still unresolved in bibliometrics.

Des exemples récents illustrent le type de problèmes que la bibliométrie s'attache à résoudre. Ils éclairent quatre sous-domaines de cette discipline: mesures de l'utilisation des ressources documentaires, analyse de citations, linguistique, modélisation des données dans des études de productivité.

Introduction

Bibliometrics is the quantitative study of interrelationships of properties of recorded discourse. Although models or equations have been proposed to describe these relationships over the past fifty years, recent economic and technological conditions in information science have presented problems for which such models do not provide satisfactory answers. The purpose of this short presentation is to describe problems which have been solved and to outline problems which have not been answered. Because of time limitations only problems from four subareas of bibliometrics will be discussed. They are: library usage management, citation analysis, linguistics, and journal and author productivity distributions.

The conditions which have forced library and information scientists to reexamine the traditional models are: 1) the increase in the amount that is published, 2) the increase in the cost of items produced, 3) the budget cut-backs in campus libraries, 4) the lack of space for current holdings, and 5) the increased use of computers in li-

brary systems. (See for example, the recent article "The Electronic Library" in the April 1983 issue of University Affairs wherein the campus library is referred to as a "data centre" and the problems associated with the advancement of computer technology are clearly explained.)

Usage Management Problems

When should a library discontinue purchasing a periodical? What variables should be used to judge when it is time to remove a journal from a shelf? How can a journal's utility be measured?

Many authors have addressed these problems with varying degrees of success. Let me briefly outline five examples.

Dhawan, Phull and Jain¹ provided an answer to these questions which was based on their experience. If libraries are considering journal acquisitions they suggest the following strategy: acquire, in the following order, those journals which are 1) used, abstracted and cited, 2) used, abstracted but not cited, 3) used, cited but not abstracted, 4) used, but neither abstracted nor cited, 5) abstracted and cited but not used. That is, find those journals which satisfy and the remaining journals are subscribed to in the order of priority given above, subject to the availability

of funds. This opinion is admirable; however, producing (ranked) lists for the hundreds of university disciplines would be costly, time-consuming and a stendous task, to say the least. I feel the solution to this problem has not been realistically solved by their suggestion.

McBain and Bobick² list 15 other articles where attempts and various strategies for collection management and storage are discussed. They also present a solution for a related localized problem: how to assess potential usage within a departmental library. Temple University has a Biology Library in one building and under direction of a main library (Paley). In 1978 they wanted to gain some information on the potential usage of its journal collection by the people in the biology department. They measured this usage by the journal citations from writings from three different user groups within the department. Table 1 summarizes these groups.

It is interesting to note that of the 24 full-time faculty only 13 had at least one publication with citations during the three years covered by this study. Each of the four rows in the table yielded a distribution relating $F(x)$, the cumulative number of citations, to x , the journal rank number (the journals were ranked in order of decreasing

* John J. Hubert is Professor at the Department of Mathematics and Statistics at the University of Guelph, Ontario.

Table 1
Group Characteristics

User Source	Number	Citations	Number of Journals Cited	Total Journal Citations
Faculty Publications	92	2047	234	1793
Ph. D. Theses	9	673	119	632
Qualifying Briefs	52	1435	196	1314
Combined	153	4155	336	3739

numbers of citations). They then ignored the first four most cited journals and all journals with only one citation. To this reduced data they fitted a Bradford³ model:

$$F(x) = a + b \log(x).$$

The patterns were quite similar and consistent with other studies. In any case their study did isolate those journals (for example pre-1960 issues) which could be removed and stored in a main library or be candidates for cancellation. Their study also emphasized the problem of obsolescence and they suggested that the development of a model of aging would have predictive value in formulating collection maintenance policies.

Miyamoto and Nakayama⁴ also addressed these problems by considering the interrelationship of four variables: utility of a journal (which was measured by the photocopy demand in a library and the citation frequency), cost (which was measured by shelf maintenance), the total number (N) of journals in the library, and t , the age of a journal. Their model for the utility function, $p_i(t)$ of the i th most used journal at time t is of the form

(1) $p_i(t) = \Delta R(i) [a_1 e^{-bt} + a_2 e^{-ct}]$ where $\Delta R(i)$ was determined from Bradford's model. This is very similar to the utility function of Brookes⁵ which states that

$$p_i(t) = (N/i) \exp(-at)$$

for its most cited journal. The main advantage of (1), they claim is that whereas the Brookes model is not everywhere optimal, their model is everywhere optimal in the sense of the "constrained optimization technique". It is an interesting application of an operations research method to isolate the best time when a journal should be removed from the collection. They do admit that they did not solve the problem of how the purchase cost of a journal could be incorporated into their utility function and it is still an open problem.

Parker⁶ addressed the problem of which books should be weeded out by proposing a model which relates $F(x)$, the cumulative number of books (or journals), to x , the number of times

it was used. Essentially the model is

$$F(x) = ax(b + x)^{-c}$$

where x where $x > 0$, and the three parameters a , b and c are positive and estimable from data. (for example, from the University of Missouri library database of 385 989 uses of 139 892 books over a 12 month period in 1972-73, he found the estimates of the parameters to be $a = 14.8$, $b = 12 100$, and $c = 3.5$.) He admits (a) that the parameters do not lend to easy interpretation, (b) that no test of adequacy of the model was made, (c) that the model overestimates and underestimates with regularity but these variations are due to "random variations" and "seasonal fluctuations" and not "from any defect of the model", and (d) no theoretical justification exists for this model. It should be noted that the literature is rich with models of this type as I will discuss later. Parker has two other models: a double exponential model for obsolescence⁷, but it is identical to the Miyamoto and Nakayama model; a hyperbolic function model⁸ for the utility function mentioned earlier which has been advocated many times. I see no new modelling contributions from these attempts.

Casper⁹ also considered the problem of modelling the demand for library service by users. This variable was measured by the number of library items requested by those who pay a user's fee at the Health Sciences Library at Case Western Reserve University in Ohio. If Y is demand, then the model was the multiple linear regression function

$$(2) \quad Y = \beta_0 + \sum \beta_j X_j + E$$

where $j = 1, 2, \dots, p=7$ regressor variables were characteristics of the hospitals to which the user was affiliated. They were:

X_1 = whether or not the hospital of the user was a library member

X_2 = total hospital expenditures for 12 months

X_3 = total payroll for 12 months

X_4 = number of persons employed

X_5 = number of beds in the hospital

X_6 = number of patients admitted per year

X_7 = mean number of inpatients receiving care per day.

Data was collected for 123 users from May 1971 to April 1982. Using the standard stepwise regression procedure the answer became

$$Y = -67.8 + 0.29X_1 - 1.84X_4 + 0.03X_6 - 0.50X_7$$

with $R^2 = 82\%$.

There are three important problems with this analysis. (a) Of the 123 sample points the top 15 (12%) accounted for 75% of the total requested and two were not members and one of these accounted for over 11% of all requests. Thus there is heterogeneity in the variability of the Y variable over the X variables. (b) Multicollinearity exists; i.e., some X variables are correlated which causes certain variables to be declared insignificant when in fact they are significant. For example, X_2 is highly correlated with X_3 (also X_3 and X_4 are highly correlated). The procedure enters only one of these expense variables, namely the one with the higher correlation with Y ; the other is set aside because it explains very little additional variation (or, it provides very little additional predictive ability to what is already given by the other variable). (c) What was measured, the number of library items requested, is best determined not only by demand, but also the supply of the library service. Casper acknowledges this: "Better estimates of the true parameters of the demand function might be obtained by estimating both demand and supply simultaneously". This problem can only be resolved by incorporating two Y values on the left hand side of (2) and then applying standard multivariate statistical procedures. This latter methodology could be applied to many of the problems posed in this presentation and I hope to see more of its use in the literature of information scientists.

Problems in Citation Analysis

How can citation counts be used as a criterion measure for assessing the impact of scientific scholarship? Is self-citation an important variable in citation analysis? How can self-citation be measured and how can citation analysis be adjusted for its influence? Can citation analysis be used to compare the productivity of different scientific disciplines?

The counting of the number of citations a journal or paper or an author

receives has generated problems and indeed controversies since the introduction of Eugene Garfield's Science Citation Index (SCI) almost 20 years ago. (See for example, the excellent articles of Smith¹⁰ and Gottfredson and Garvey.)¹¹

In the following paragraphs I will report on some examples which attempt to answer the above problems.

One of the problems in citation analysis is the concept of self-citation. It is the principal source of objections to the citation analysis itself. It has been called "scholarly bricklaying" and its use has been called "ego-enhancement".

Lawani¹² dissected the concept of citation and presented a systematic classification of the types of citation, especially self-citation. Is it an important issue? There is individual self-citation. (Lawani reports that the self-citation rates are 25% for authors in cancer research, it is 16% for plant physiologists and it is 18% for neurobiologists.) There is subject field to subject field self-citation (50-60% in the social sciences and 70-80% in science and technology) and there is journal to journal citation (the rates are 52% for the *Agronomy Journal*, 60% for *Crop Science*, 83% for *Annales Agronomiques* and 100% for *L'Agonomie Tropicale*). Lawani also asks: why do we have this phenomenon? Are the rates increasing? What are the sociological factors generating this phenomenon? Is it a need for research collaboration or is it egotism? How can we measure the influence of these variables? I would agree that these are questions that need to be answered.

Lawani and Bayer¹³ addressed the question of how citation counts can be used as a criterion measure for assessing scientific scholarship. Their paper is an excellent example of a valid research design and legitimate methodology. Their basic question was: what is the relationship between the citation counts of a scientific community for a set of papers and the quality of these papers as judged by peers. They obtained a random sample of 870 articles in cancer research and without going into detail other than to say that regression analysis plays a role and self-citations were accounted for, they discovered that the two variables, citation rates and peer assessment, were highly correlated. That is, both variables are reasonably reliable and largely interchangeable as indices for evaluation within the set of authors

considered. However, neither variable can be reliably used to differentiate between individuals or between their scientific contributions. For example, they pose the question of why there are still unrated but highly cited papers as well as highly rated but poorly cited papers. The methodology is valid for group comparisons but not meaningful for individual comparisons.

Can citation analysis play a role in stock management in a library? I believe the examples of McBain and Bobick² and Miyamoto and Nakayama³ cited earlier in this presentation have shown that these counts can be used for such problems.

Do the patterns observed in frequency tables summarizing the citation counts of groups of individuals have any regularity? Can they be modelled?

The study of the frequency distributions of citation counts by physicists has been an ongoing study by Vlady¹⁴ for 10 years. He is convinced that the model originally proposed by Lotka^{15,16} adequately described the observed patterns.

Is there a difference between the behaviour of citation frequency distributions of chemists and physicists? Do disciplines have their own characteristic citation pattern?

Dick et al.¹⁷ recently studied the two year citation counts of three different groups of physical scientists and discovered that (a) the patterns can be adequately described by a rank-frequency model of the form $f(r) = ae^{-br}$, where $f(r)$ is the number of citations of a scientist of rank r and a and b are positive-valued parameters. (The authors were ranked by decreasing order of counts and ties were assigned the maximal rank.) On log-log paper the patterns became negative-sloping straight lines. The slopes were significantly different and b , the slope parameter, apparently characterized the disciplines.

Do these patterns remain stable over time? Do chemists and mathematicians have different citing habits?

Problems in Linguistics

Linguists have recognized that language can be spoken, thought, written, and expressed in shorthand, Morse code, Braille, teleprinter code, and even whistled. Each of these 'codings' has been analyzed statistically in numerous ways but the most common approaches are: counting the number

of occurrences of words, counting the number of times a word of given length occurs, counting the number of times a sentence of given length occurs, and counting the number of times a word containing a given number of phonemes occurs. The resulting frequency distributions have been used to characterize an author or indeed a language. They have been used to estimate an author's total vocabulary and to decide on a text's authorship.

The science of statistical linguistics has a long history and the literature is rich with examples and models. A comprehensive reference list can be constructed from the lists in ref. 18 to 23.

The following is a single recent example of the type of linguistic problems which have received renewed interest. Given that there is a relationship between word length and its frequency of occurrence, are there methods which will permit meaningful comparisons within and among languages? Many authors have developed algorithms and models to distinguish writers and languages. But are there situations where there is some regularity?

This was the subject of the excellent article by Parunak²⁴. She has developed a graphical technique which can filter out patterns and suggest models. One consequence of her study was that she found that a form of the incomplete gamma function models certain relationships. When this technique was applied to several large texts in Greek, French and English, she discovered that as the word length changes, the patterns change in a remarkably regular fashion. This regularity was independent of language and text length. However, when the data is grouped on certain factors such as first letters of words, the patterns become highly irregular. The important point of this example is that her results suggest that this technique could be useful in analyzing rank-size distributions found so abundantly in various other bibliometric situations.

Problems in Modelling Productivity Distributions

Almost all bibliometric models relate, in a simple functional form, one variable with another variable. For example, in journal productivity studies, for a bibliography covering a certain span of years on a particular subject, a few journals contribute a large number of articles, other journals contribute fewer, and so on in a monotonic se-

quence ending with a large number of journals contributing one article each to the subject. The two variables are number of journals and number of articles. After arranging the journals in a decreasing order of productivity, a size-frequency distribution is obtained for the number of journals containing a fixed number of articles each. Conversely, a rank-frequency table can be constructed for the number of articles associated with a journal of fixed rank. These two approaches to observed patterns form the two modes of the data tabulations.

Over the past fifty years, a sizable body of literature dealing with bibliometric models has developed. The early models were proposed because they were observed to fit graphically certain specific empirical frequency distributions. In many cases their functional forms were identical, the similarity only noted by other writers years later. In each case, depending on the subject field they applied to, there was a proliferation of papers which modified, extended, clarified, applied, and generalized the initial model.

To illustrate explicitly the notions of the size-frequency approach, consider the data in Table 2 which has been adapted from Bradford³. $f(n)$ denotes the number of journals contributing exactly n articles each to a particular subject field such that the total number of observed journals is $J = \sum f(n)$ and the total number of observed articles is $N = \sum nf(n)$. This tabulation relates the observations (the articles) with a class (a journal). The rank-frequency approach inverts this tabulation. You will notice that near the bottom of the table there is one journal contributing the most articles (22). This journal is assigned the rank 1. The next most productive journal is assigned rank 2 because it contributed 18 articles. This is continued, resulting in the rank-frequency distribution given on the right side of Table 2, where $g(r)$ is the number of articles contributed by the journal of rank r . Notice that there are two journals contributing 13 papers each, and each is assigned rank 5, the "maximal-rank" assignment method which is used in the case of ties. (If we assign the rank 4 of each of these journals, then we are using a "minimum-rank" method; there are also the random-rank and average-rank methods.) The rank-frequency tabulation reverses the order of the size-frequency tabulation, and gives priority to the most productive journals. The size-

Table 2
The Size-frequency Distribution of the Number of Journals ($f(n)$) Contributing n Papers Each and the Rank-frequency Distribution of the Number of Papers $g(r)$ Contributed by a Journal of Rank r

n	$f(n)$	$nf(n)$	r	$g(r)$
1	102	102	1	22
2	25	50	2	18
3	13	39	3	15
4	2	8	5	13
5	7	35	7	10
6	1	6	8	9
7	3	21	11	9
8	3	24	14	7
9	1	9	15	6
10	2	20	22	5
13	2	26	24	4
15	1	15	37	3
18	1	18	62	2
22	1	22	64	1
Sum	$J=164$	$N=395$		

frequency approach gives emphasis to the journals of least productivity. There are other relationships between the two approaches. Properties of these approaches are discussed in ref. 25 to 32.

The modelling problem is to find a mathematical equation relating $f(n)$ with n . Associated problems are: What is the process which generates this relationship? What happens to the relationship if a larger sample of observations, N , is obtained? Does the relationship remain the same from year to year? What are the statistical interrelationships between the various univariate models claiming to describe these relationships?

In the last 25 years, it has been observed that such tabulations occur for other pairs of variables from a wide variety of natural and social phenomena. Table 3 provides some examples of such combinations of observation versus class relationship.

Table 6
Examples of Observation-class Relationship

Observation	Class
Number of articles	journals
Number of papers	authors
Number of citations	persons
Number of occurrences	nouns
Length of word	words
Checked-out frequency	books
Length of sentence	sentences
Number of phonemes	words
Number of insects	species
Number of occurrences	initial digits
Income level	persons
Size of cities	cities

A list of all the models that have been proposed in the literature would have to include the names of Lotka, Yule, Pareto, Zipf, Bradford, Vickery, Mandelbrot, Hardan, Good, Simon, Leimkuhler, Cole, Naranan, Hill, Price, and Brookes. Besides the many papers showing the applicability of these variations, some recent papers are mentioned in notes.³³⁻⁴⁰ Also, thorough discussions of the mathematical aspects of some of these models can be found in references 41 to 43.

The questions asked earlier have been answered more than once by the references indicated in the last paragraph. I see no purpose to adding still more models or so-called general forms of this univariate relationship. Personally I feel the phenomenon may be more complex. What is now the problem is to extend the empirical studies to higher order models and to multivariate models which can incorporate more information into the model including the interaction effects. The use of these advanced techniques can become the next exciting era of bibliometric modelling.

1. Dhawan, S.M., Phull, S.K., and Jain, S.P., «Selection of Scientific Journals: a Model», *Journal of Documentation*, vol. 36, n° 1 (1980), pp. 24-31.

2. McCain, K.W. and Bobick, J.E., «Patterns of Journal Use in a Departmental Library: a Citation Analysis», *Journal of the American Society for Information Science*, vol. 32 (1981), pp. 257-267.

3. Bradford, S.C., «Sources of Information on Specific Subjects», *Engineering*, vol. 137 (1934), pp. 85-86.

4. Miyamoto, S. and Nakayama, K., «Determination of the Conservation Time of Periodi-

- cals for Optimal Shelf Maintenance of a Library», *Journal of the American Society for Information Science*, vol. 32 (1981), pp. 268-274.
5. Brookes, B.C., «Optimum P% Library of Scientific Periodicals», *Nature*, vol. 232, n° 13 (1971), pp. 458-461.
6. Parker, R.H., «Bibliometric Models for Management of an Information Store. III. Developing an Empirical Model», *Journal of the American Society for Information Science*, vol. 33 (1982), pp. 134-138.
7. Parker, R.H., «Bibliometric Models for Management of an Information Store. II. Use as a Function of Age of Materials», *Journal of the American Society for Information Science*, vol. 33 (1982), pp. 129-133.
8. Parker, R.H., «Bibliometric Models for Management of an Information Store. I. Differential Utility among Items», *Journal of the American Society for Information Science*, vol. 33 (1982), pp. 124-128.
9. Casper, C.A., «Estimating the Demand for Library Service: Theory and Practice», *Journal of the American Society for Information Science*, vol. 29 (1978), pp. 232-237.
10. Smith, L.C., «Citation Analysis», *Library Trends*, No. 2, (Summer 1981).
11. Gottfredson, S.D. and Garvey, W.D., «Citation Indexing: its Theory and Application in Science, Technology, and Humanities» (Review), *Behavioral and Social Science Librarian*, vol. 1 (1980), pp. 289-294.
12. Lawani, S.M., «On the Heterogeneity and Classification of Author Self-Citations», *Journal of the American Society for Information Science*, vol. 33, n° 3 (1982), pp. 281-284.
13. Lawani, S.M. and Bayer, A.E., «Validity of Citation Criteria for Assessing the Influence of Scientific Publications: New Evidence with Peer Assessment», *Journal of the American Society for Information Science*, vol. 34, n° 1 (1983), pp. 59-66.
14. Vlachy, J., «Citation Analysis in Particle Physics», *Czechoslovakian Journal of Physics, Series B*, vol. 32 (1982), pp. 1182-1134.
15. Lotka, A.J., «The Frequency Distribution of Scientific Productivity», *Journal of the Washington Academy of Sciences*, vol. 16 (1926), pp. 317-323.
16. Hubert, J.J., «Lotka's Law in the Humanities», *Journal of the American Society for Information Science*, vol. 28 (1978), p. 66.
17. Dick, J.P., Hubert, J.J., and Smith, W.G., «A Model for Citation Counts for Physical Scientists» (To appear in *Scientometrics*).
18. Good, I.J., «Distribution of Word Frequencies», *Nature*, n° 179 (1957), p. 595.
19. McNeil, D.R., «Estimating an Author's Vocabulary», *Journal of the American Statistical Association*, vol. 68 (1973), pp. 92-96.
20. Gani, J. and Saunders, I., «Some Vocabulary Studies of Literary Texts», *Sankhya, Series B*, vol. 38 (1976), pp. 101-111.
21. Brainerd, B., *Weighting Evidence in Language and Literature: A Statistical Approach* (University of Toronto Press, Toronto, 1974).
22. Hubert, J.J., «Linguistic Indicators», *Social Indicators Research Journal*, vol. 8 (1980), pp. 223-255.
23. Tuchinsky, P., «Zipf's Law and His Efforts to Use Infinite Series in Linguistics», *Journal of Undergraduate Mathematics and Its Applications, Mathematical Association of America, Module 213* (1980), pp. 80-100.
24. Parunak, A., «Graphical Analysis of Ranked Counts (of Words)», *Journal of the American Statistical Association*, vol. 74, n° 365 (1979), pp. 25-30.
25. Hubert, J.J., «Bibliometric Models for Journal Productivity», *Social Indicators Research Journal*, vol. 4 (1977), pp. 441-473.
26. Hubert, J.J., «A Relationship Between Two Forms of Bradford's Law», *Journal of the American Society for Information Science*, vol. 29 (1978), pp. 159-161.
27. Brookes, B.C. and Griffiths, J.M., «Frequency-Rank Distribution», *Journal of the American Society for Information Science*, vol. 29 (1978), pp. 5-13.
28. Hubert, J.J., «General Bibliometric Models», *Library Trends* (Summer 1981), pp. 65-81.
29. Fairthorne, R.A., «Empirical Hyperbolic Distributions (Bradford-Zipf-Mandelbrot) for Bibliometric Description and Prediction», *Journal of Documentation*, vol. 25 (1969), pp. 319-343.
30. Bookstein, A., «The Bibliometric Distributions», *Library Quarterly*, vol. 46 (1976), pp. 416-423.
31. Brookes, B.C., «Theory of the Bradford Laws», *Journal of Documentation*, vol. 33 (1977), pp. 180-209.
32. Bookstein, A., «Explanations of the Bibliometric Laws», *Collection Management*, vol. 3 (1979), pp. 151-162.
33. Asai, I., «A General Formulation of Bradford's Distribution: the Graph-Oriented Approach», *Journal of the American Society for Information Science*, vol. 32, n° 2 (1981), pp. 113-119.
34. Asai, I., «Adjusted Age Distribution and its Application to Impact Factor and Immediacy Index», *Journal of the American Society for Information Science*, vol. 32, n° 3 (1981), pp. 172-174.
35. Bulick, S., «Book Use as a Bradford-Zipf Phenomenon», *College and Research Libraries*, vol. 39, n° 3 (1978), pp. 215-219.
36. Lawani, S.M., «Citation Analysis and the Quality of Scientific Productivity», *Bioscience*, vol. 27 (1977), pp. 26-31.
37. Burrell, Q., «A Simple Stochastic Model for Library Loans», *Journal of Documentation*, vol. 36, n° 2 (1978), pp. 115-132.
38. Burrell, Q., «Alternative Models for Library Circulation Data», *Journal of Documentation*, vol. 38, n° 1 (1982), pp. 1-13.
39. Hindle, A. and Worthington, D., «Simple Stochastic Models for Library Loans», *Journal of Documentation*, vol. 36, n° 3 (1980), pp. 209-213.
40. Hubert, J.J., «A Rank-Frequency Model for Scientific Productivity», *Scientometrics*, vol. 3, n° 3 (1981), pp. 191-202.
41. Brookes, B.C., «A Critical Commentary on Leimkuhler's 'Exact' Formulation of the Bradford Laws», *Journal of Documentation*, vol. 37, n° 2 (1981), pp. 77-88.
42. Fedorowicz, J., «The Theoretical Foundation of Zipf's Law and its Application to the Bibliographic Database Environment», *Journal of the American Society for Information Science*, vol. 33, n° 3 (1982), pp. 285-293.
43. Tague, J., «The Success-Breeds-Success Phenomenon and Bibliometric Processes», *Journal of the American Society for Information Science*, vol. 32 (1981), pp. 280-286.

Écrivez-nous pour recevoir notre nouveau catalogue d'ouvrages en informatique



McGraw-Hill, Éditeurs
750, boul. Laurentien, SAINT-LAURENT H4M 2M4
téléphone: (514) 744-5531

NOUVELLE PUBLICATION DANS LA SÉRIE

BIBLIOGRAPHIE DU SPORT



TOMES 1 ET 2

MISE À JOUR

Un guide qui vous initiera à la littérature sportive publiée de 1980 à 1982.

La **MISE À JOUR** trouvera tout naturellement sa place sur les étagères des bibliothèques de quiconque possède déjà la **BIBLIOGRAPHIE DU SPORT**. Les deux tomes réunissent plus de 30 000 références bibliographiques supplémentaires.

MISE À JOUR DE LA BIBLIOGRAPHIE DU SPORT TOME 1

Mise à jour des volumes 1 à 4 de la **BIBLIOGRAPHIE DU SPORT** sur le sport et les activités sportives.

MISE À JOUR DE LA BIBLIOGRAPHIE DU SPORT TOME 2

Mise à jour des volumes 5 à 8 de la **BIBLIOGRAPHIE DU SPORT** qui traitent de tous les sports, de la médecine sportive, de l'histoire, de l'entraînement, de la psychologie sportive et des compétitions.



BIBLIOGRAPHIE DU SPORT dont les 1 600 pages fournissent

plus de 70 000 références bibliographiques.

l'index des sujets le plus complet jamais publié dans le domaine du sport!

publié par le

**CENTRE DE DOCUMENTATION
POUR LE SPORT**



Un service de
L'ASSOCIATION CANADIENNE
DES ENTRAÎNEURS

SI VOUS DÉSIREZ DES RENSEIGNEMENTS SUPPLÉMENTAIRES CONCERNANT CES VOLUMES OU SI VOUS DÉSIREZ LES COMMANDER, COMMUNIQUEZ AVEC LE:

CENTRE DE DOCUMENTATION
POUR LE SPORT
333 RIVER ROAD
OTTAWA K1L 8H9
ONTARIO

PRIX

MISE À JOUR DE LA BIBLIOGRAPHIE DU SPORT

TOME 1 — \$50

TOME 2 — \$50

Commandez les deux pour la modique somme de \$90. Vous économiserez ainsi \$10.

BIBLIOGRAPHIE DU SPORT

Série de huit volumes \$175

Chaque volume \$25 l'unité

Veuillez ajouter 1.50 par volume pour les frais d'envoi et de manutention — 3 dollars à l'extérieur du Canada

LES RÉSIDENTS DE L'ONTARIO SONT PRIÉS D'AJOUTER 7% DE TAXE SUR LA VENTE AU COÛT TOTAL DE LA COMMANDE.

Recherche sur l'interface système-usager des services de télé-référence

Par Gilles Deschatelets*

L'exposé présente les principaux résultats d'un projet de recherche sur le rôle et l'importance des divers intervenants dans la recherche bibliographique en-ligne (télé-référence), et qui voulait répondre aux questions suivantes: 1) Existe-t-il un mode d'interface optimal qui pourrait être généralisé en télé-référence? 2) Le mode d'interface est-il sensible à certaines caractéristiques de l'utilisateur et de sa question de recherche? 3) Si oui, pouvons-nous, à partir de ces caractéristiques, prédire un mode d'interface qui maximiserait une ou plusieurs des mesures de performance? Plus d'une cinquantaine de variables ont été identifiées, mesurées et analysées et neuf hypothèses de recherche ont été vérifiées. Enfin, une analyse fut réalisée dans le but de déterminer une fonction statistique permettant de prédire le mode d'interface optimisant diverses mesures de performance.

The article presents the major results of a research project on the role and impact of various intermediaries during an on-line bibliographic search. The goal of the study was to reply to the following questions: 1) Is there an optimal interface which can be generalized to on-line searching? 2) Is this interface sensitive to certain user characteristics and search questions? 3) If so, on the basis of these characteristics, can we predict an interface which would maximize one or several performance criteria? Over fifty variables were identified, measured and analysed as well as nine research hypothesis verified. Finally, an analysis was carried out to determine a statistical function related to the interface and performance criteria.

Introduction

L'exploitation systématique de l'ordinateur pour la compilation de bibliographies représente très certainement l'un des principaux événements ayant marqué l'évolution de la bibliothéconomie et des sciences de l'information. Avec un minimum de temps et d'effort, l'ordinateur permet la compilation de bibliographies et la réalisation de recherches documentaires complexes, exhaustives et surtout, personnalisées. La télé-référence, c'est-à-dire la compilation de bibliographies par télématique permet au chercheur, au professeur, à l'étudiant, au décideur, à l'administrateur et à tout autre utilisateur, de satisfaire rapidement ses besoins spécifiques de données ou d'information documentaire.

Ce faisant, l'ordinateur nous fait passer de l'ère de la diffusion de type

encyclopédique à celle de la diffusion et du repérage sélectif ou individuel de l'information. En général, cette nouvelle procédure informatique ne diffère pas tellement de la bonne vieille bibliographie «manuelle». De fait, la démarche intellectuelle des deux méthodes est exactement la même.

Nous pouvons signaler toutefois deux distinctions majeures entre les deux procédures: 1) dans la compilation manuelle, la démarche intellectuelle, la stratégie de la recherche est la plupart du temps implicite tandis que dans la compilation informatisée, toutes les étapes et tous les éléments de la recherche doivent être clairement identifiés et spécifiés à l'ordinateur; 2) dans la compilation manuelle, le dépouillement de répertoires imprimés n'implique, comme médium de communication avec le message imprimé, que le processus de la lecture tandis qu'à cela s'ajoute, dans la compilation informatisée, l'indispensable connaissance d'un «langage» pour dialoguer avec l'ordinateur.

Pour ces deux raisons au moins, la télé-référence est demeurée, jusqu'à présent, une sorte de chasse gardée des médiateurs de recherche, c'est-à-dire, pour la plupart, des bibliothécaires et documentalistes, qui apprennent à maîtriser parfaitement tous les aspects techniques de la recherche et qui, en outre, de par leur formation, peuvent aider l'utilisateur à expliciter sa stratégie de recherche et à «formuler» sa question selon les exigences du système.

En fait, lorsqu'on analyse un tant soit peu la littérature sur le sujet, on se rend compte que s'est développée, avec l'utilisation des systèmes en-ligne pour la recherche documentaire, une véritable problématique de l'interface homme-machine. D'un côté, les bibliothécaires et documentalistes prétendent que la fonction de médiateur leur revient de droit puisqu'ils jouent déjà ce rôle auprès de leurs clientèles. Seul le support technologique a changé. Parmi les arguments invoqués à l'appui de cette prétention, mentionnons:

* Gilles Deschatelets est directeur de la Bibliothèque scientifique de l'Université Laval à Québec.

- 1° – la qualité supérieure de leurs recherches;
- 2° – un bien meilleur taux coût-efficacité de leurs recherches;
- 3° – leur meilleure connaissance de la variété, du contenu et de la structure des banques d'information;
- 4° – leur meilleure connaissance de la variété, du contenu et des langages des systèmes d'information;
- 5° – un meilleur suivi des innombrables changements aux banques et aux systèmes;
- 6° – les coûts prohibitifs de formation des usagers;
- 7° – le trop peu d'utilisation des systèmes par les usagers pour que ces derniers maintiennent leur expertise à un niveau satisfaisant; et enfin
- 8° – un dernier argument qu'on pourrait appeler la loi de Skippy ou le «syndrome de la tartine de beurre d'arachides.» En fait, c'est une variante de la loi de Murphy qui dit que si vous échappez une tartine de beurre d'arachides par terre, elle va probablement tomber du côté beurré. De même, tout ce qui est susceptible de mal aller pendant une recherche au terminal ira effectivement très mal. D'où la nécessité d'une expertise humaine pendant la recherche. Or, ce n'est pas l'utilisateur mais le médiateur qui possède cette expertise.

Voilà pour les arguments en faveur des recherches médiatisées. De l'autre côté, se retrouvent les défenseurs des recherches non-médiatisées, dont les principaux arguments sont:

- 1° – que le médiateur accroît considérablement le coût des recherches;
- 2° – qu'il existe de nombreux dangers et problèmes de communication de la question et de son interprétation par le médiateur;
- 3° – que l'aspect heuristique de la recherche est fort important, de même que le «bouquinage», magasinage que l'on peut faire pendant une recherche;
- 4° – qu'un minimum de formation est nécessaire pour être capable d'effectuer une recherche au terminal; et enfin
- 5° – le syndrome de la «tartine de beurre d'arachides... beurrée

des deux côtés», c'est-à-dire qui doit nécessairement tomber sur un côté beurré; en d'autres termes, ils prétendent qu'il y a un réel danger de créer volontairement des systèmes tellement compliqués à utiliser qu'ils assurent un «emploi pour les amis», c'est-à-dire les médiateurs.

Enfin, entre ces deux extrêmes, se retrouvent les sages, les pondérés, qui proposent de joindre l'utile à l'agréable, le médiateur et le médicé, et faire effectuer les recherches au terminal par le médiateur initié en présence de l'utilisateur érudit. Ce faisant, on évite les défauts des deux autres modes d'interrogation pour n'en conserver que les avantages. Autrement dit, la tartine tombe sur la croûte et reste en équilibre...

Voilà, très grossièrement défini, le problème auquel s'adressait la recherche, effectuée de 1979 à 1982. Le but visé était d'analyser et de comparer divers modes d'interface entre l'utilisateur et le système, le mode d'interface étant défini, pour les besoins de l'étude, comme un mode d'interrogation du système. De façon opérationnelle, trois modes d'interrogation ont donc été retenus pour fins d'analyse et de comparaison: a) *le mode délégué*, dans lequel le médiateur effectue seul la recherche pour l'utilisateur, b) *le mode conjoint*, dans lequel le médiateur et l'utilisateur effectuent la recherche ensemble au terminal et c) *le mode direct*, dans lequel l'utilisateur effectue lui-même sa recherche sans l'aide du médiateur.

La recherche avait pour but de répondre à trois questions:

- 1) existe-t-il selon un mode d'interface optimal (c'est-à-dire supérieur selon diverses mesures de performance) qui pourrait être généralisé en téléconsultation?
- 2) le mode d'interface est-il sensible à certaines caractéristiques de l'utilisateur et de sa question de recherche?
- 3) si oui, est-il possible, à partir de ces caractéristiques, de prédire pour un nouvel utilisateur ou pour une population d'utilisateurs, un mode d'interface optimal qui maximiserait une ou plusieurs mesures de performance?

Notre exposé s'adresse exclusivement à la première question. Les deux autres ont été abordées au Congrès de l'Association canadienne des sciences de l'information en mai 1983, à Halifax.

L'objectif poursuivi était donc d'analyser en profondeur cette question de l'interface entre l'utilisateur et le système, dans une optique de gestion et de performance, afin d'en démontrer l'importance et le coût aux centres de service de téléconsultation et aux institutions qui les chapeautent et qui, bien souvent, les financent.

Depuis les débuts du projet, en mars 1979, l'infrastructure technologique des services de téléconsultation a considérablement évolué. De nombreuses recherches ont été entreprises, dans les dernières années, notamment pour améliorer et faciliter le dialogue homme-machine et pour rendre presque naturel le langage d'interrogation des systèmes d'information. Le modèle traditionnel du chercheur assis à son terminal interrogeant, une à une, chacune des banques nécessaires à sa recherche documentaire, répétant à chaque fois sa stratégie de recherche, maîtrisant parfaitement le langage de chaque système utilisé et les procédures de communications pour s'y rendre, ce modèle traditionnel, dis-je, est déjà en bonne voie de devenir anachronique. Soit dit en passant, n'est-ce pas étrange de parler de modèle «traditionnel» pour des systèmes qui datent d'à peine dix ans? La sophistication rapide des logiciels d'interrogation et l'essor fantastique de la micro-informatique, notamment, permettent déjà – ou dans certains cas le permettront dans un avenir rapproché – la consultation simultanée de plusieurs banques, la mise en mémoire des recherches, la traduction automatique des logiciels d'interrogation en un seul langage de son choix et toute une gamme de programmes d'aide à l'utilisateur selon son niveau d'expérience et d'expertise.

Loin de contredire ou d'invalider les résultats de la présente étude, tous ces développements technologiques ne contribueront à notre avis qu'à accentuer davantage les caractéristiques de chaque mode d'interrogation et leurs différences fondamentales.

Mesures de performance

Neuf mesures de performance furent choisies pour analyser et comparer les trois modes d'interrogation (voir le tableau 1).

Efficacité de la recherche (SE)

L'efficacité de la recherche est définie comme le nombre de références pertinentes imprimées sur le nombre total de références examinées. C'est une me-

sure familière en sciences de l'information pour évaluer l'habileté du chercheur qui interroge des systèmes d'information, à ne repérer que des références pertinentes. On y réfère souvent sous l'expression «taux de précision». Cette mesure avait d'abord été utilisée avec les systèmes de repérage en mode différé. En réalité, elle n'est pas vraiment tout à fait adéquate en télé référence parce que l'interrogation en mode conversationnel est une activité heuristique et que le nombre de références, et de références pertinentes, à imprimer est sous le contrôle immédiat du chercheur qui peut porter ses jugements de pertinence immédiatement pendant la recherche. Le taux de précision demeure toutefois fort acceptable comme critère de comparaison des modes d'interrogation. En effet, il apporte un certain éclairage sur le comportement du chercheur, c'est-à-dire sur le degré «d'heuristique» démontré dans chaque mode d'interrogation.

Dans cette optique, le taux de précision utilisé dans la présente étude fut défini un peu différemment de sa connotation traditionnelle, c'est-à-dire nombre de références pertinentes sur nombre total de références repérées. Le

terme «références repérées» fut explicité davantage par «références examinées» et «références imprimées». Examiner une référence réfère à l'impression en ligne d'assez d'information sur une référence pour rendre capable de porter un jugement de pertinence. «Assez d'information», cela peut aller de l'impression du titre seul jusqu'à l'impression du dossier complet de la référence. Imprimer une référence, en direct ou en différé, d'autre part, réfère à l'impression d'assez d'information sur une référence pour rendre capable d'identifier et de localiser le document décrit. En télé référence, un premier examen des titres de documents pour effectuer ensuite le choix des références à imprimer est devenu pratique courante.

Efficacité du mode d'interrogation (SI) et Efficacité du chercheur (SGE)

L'efficacité du mode d'interrogation est ici associée à l'efficacité de la recherche comme mesure de performance du repérage de l'information. Ordinairement, l'efficacité de la recherche est définie comme le temps moyen de recherche nécessaire pour repérer chaque référence pertinente.

Comme l'objectif principal de la présente étude n'était pas uniquement de comparer la performance des chercheurs mais surtout celle des modes d'interrogation, l'efficacité du mode d'interrogation est définie comme «le temps total de la recherche sur le nombre de références pertinentes imprimées». Le temps total inclut, pour chacun des intervenants (usager, médiateur) s'il y a lieu, le temps de préparation, de l'entrevue préliminaire, de la recherche au terminal et de l'entrevue subséquente (voir Tableau 2).

Quant à l'efficacité du chercheur, elle a également été définie en gardant à l'esprit que cette mesure allait fort probablement favoriser les médiateurs au détriment des usagers, nettement moins expérimentés dans l'ensemble. L'efficacité du chercheur est définie comme le temps de connexion sur le nombre de références pertinentes imprimées, donc le temps de connexion ou d'interrogation moyen nécessaire pour repérer chaque référence pertinente.

Coût-efficacité du mode d'interrogation (CE)

Le coût-efficacité d'un mode d'interrogation est l'équivalent de son taux

Nom	Symbole	Définition	Mesure
Efficacité de la recherche (ou taux de précision)	SE	Nombre de références pertinentes imprimées Nombre total de références examinées	%
Efficacité du mode d'interrogation	SI	Temps total de la recherche Nombre de références pertinentes imprimées	minutes
Efficacité du chercheur	SGE	Temps de connexion de la recherche Nombre de références pertinentes imprimées	minutes
Coût-efficacité du mode d'interrogation	CE	Coût total de la recherche Nombre de références pertinentes imprimées	\$
Coût-efficacité du chercheur	CI	Coût de connexion de la recherche Nombre de références pertinentes imprimées	\$
Effort total requis dans chaque mode d'interrogation	SEF	Coût de préparation de la recherche Nombre de références pertinentes imprimées	\$
Effort requis de l'usager	USEF	Temps total consacré à la recherche par l'usager Nombre de références pertinentes imprimées	minutes
Satisfaction de l'usager	US	$\frac{US_N + US_P + US_M \times US_I + US^T}{30}$	%
Utilité de la recherche	SU	Nombre de documents à consulter Nombre de références pertinentes imprimées	%

Tableau 1. Mesures de performance: définitions

- * US_N = Satisfaction de l'usager du nombre de références repérées (de 0 à 6)
- US_P = Satisfaction de l'usager de la pertinence des références repérées (de 0 à 6)
- US_M = Satisfaction de l'usager du mode d'interrogation (de 0 à 6)

- US_I = Satisfaction de l'usager de l'utilité de la recherche (de 0 à 6)
- US^T = Satisfaction de l'usager de la télé référence en général (de 0 à 6)

Tableau 2: Calcul des temps de recherche (en minutes)

$$\begin{aligned} \text{TDIR} &= P_u + R_u \\ \text{TDEL} &= \sum P_{um} + (E_u + E_m) \\ &\quad + R_m + (A_u + A_m) \\ \text{TCONJ} &= \sum P_u + (E_u + E_m) \\ &\quad + (R_u + R_m) + (A_u \\ &\quad + A_m) \end{aligned}$$

où

$$\begin{aligned} \text{TDIR} &= \text{Temps total d'une recherche en mode direct} \\ \text{TDEL} &= \text{Temps total d'une recherche en mode délégué} \\ \text{TCONJ} &= \text{Temps total d'une recherche un mode conjoint} \\ P_u &= \text{Temps de préparation de l'utilisateur} \\ R_u &= \text{Temps de recherche de l'utilisateur} \\ E_u &= \text{Temps de l'utilisateur pour l'entrevue préliminaire} \\ A_u &= \text{Temps de l'utilisateur pour l'entrevue post-recherche} \\ P_{um} &= \text{Temps de préparation calculé globalement pour l'utilisateur et le médiateur} \\ R_m &= \text{Temps de recherche pour le médiateur} \\ E_m &= \text{Temps du médiateur pour l'entrevue préliminaire} \\ A_m &= \text{Temps du médiateur pour l'entrevue post-recherche} \end{aligned}$$

d'efficacité exprimé en dollars. On le calcule en convertissant toutes les mesures de temps en dollars, en utilisant le salaire horaire de chaque intervenant et en y ajoutant tous les coûts de la recherche elle-même (coûts de connexion, de communication et d'impression) (voir Tableau 3).

Coût-efficacité du chercheur (CI)

De la même façon, le coût-efficacité du chercheur est l'équivalent du taux d'efficacité du chercheur exprimé en dollars. Il mesure l'habileté de chaque chercheur à repérer chaque référence pertinente aussi économiquement que possible.

L'effort requis dans chaque mode d'interrogation (SEF) et l'effort requis de l'utilisateur (USEF)

Tableau 3: Calcul des coûts de recherche

$$\begin{aligned} \text{CDIR} &= T_u + C_r \\ \text{CDEL} &= (T_u + T_m) + C_r \\ \text{CCONJ} &= (T_u + T_m) + C_r \end{aligned}$$

où

$$\begin{aligned} \text{CDIR} &= \text{Coût total d'une recherche en mode direct} \\ \text{CDEL} &= \text{Coût total d'une recherche en mode délégué} \\ \text{CCONJ} &= \text{Coût total d'une recherche en mode conjoint} \\ T_u &= \text{Temps-salaire de l'utilisateur} \\ &\quad \text{TDIR} \times \text{salaire horaire de l'utilisateur} \\ &\quad \text{TDEL} (P_u + E_u + A_u) \times \text{salaire horaire de l'utilisateur} \\ &\quad \text{TCONJ} (P_u + E_u + R_u + A_u) \times \text{salaire horaire de l'utilisateur} \\ T_m &= \text{Temps-salaire du médiateur} \\ &\quad \text{TDIR} = 0 \\ &\quad \text{TDEL} = (P_m + E_m + R_m + A_m) \times \text{salaire horaire du médiateur} \\ &\quad \text{TCONJ} (P_m + E_m + R_m + A_m) \times \text{salaire horaire du médiateur} \\ C_r &= \text{Coût de la recherche} \\ &= C_{cn} + C_{cm} + C_i \end{aligned}$$

où

$$\begin{aligned} C_{cn} &= \text{coût de connexion} \\ C_{cm} &= \text{coût des communications} \\ C_i &= \text{coût de l'impression} \end{aligned}$$

L'effort requis dans chaque mode d'interrogation est défini comme le coût total de préparation sur le nombre de références pertinentes imprimées. Le coût de préparation inclut, s'il y a lieu, le coût de préparation de l'utilisateur et du médiateur et le coût de l'entrevue préliminaire. C'est donc, exprimé en dollars, l'effort total de préparation nécessaire pour repérer chaque référence pertinente.

D'autre part, *l'effort requis de l'utilisateur* est défini comme le temps total investi par l'utilisateur dans sa recherche sur le

nombre de références pertinentes imprimées. Le temps total investi par l'utilisateur inclut, s'il y a lieu, le temps de préparation, de l'entrevue préliminaire, de la recherche au terminal et de l'entrevue subséquente.

Satisfaction de l'utilisateur (US) et utilité de la recherche (SU)

Nous pouvons qualifier ces deux dernières mesures davantage comme des mesures de bénéfice de la recherche que comme des mesures de performance.

La satisfaction de l'utilisateur est la mesure du contentement ou de l'appréciation de l'utilisateur pour sa recherche en général et pour divers aspects de celle-ci. Dans la présente étude, la satisfaction de l'utilisateur est définie comme le pourcentage suivant:

la somme totale des scores de satisfaction (mesurés sur une échelle de 0 à 6) de l'utilisateur comprenant le nombre total de références repérées + la pertinence des références repérées + le mode d'interrogation + l'utilité de la recherche + la télé-référence en général, sur 30, soit le score maximum possible.

Enfin, *l'utilité de la recherche* est une mesure de la capacité de la recherche à satisfaire son ou ses objectifs. En principe, cette mesure devrait être définie comme le nombre de documents pertinents consultés sur le nombre de références pertinentes imprimées. Or à cause de restrictions de temps, il nous était impossible de calculer le nombre de documents pertinents consultés, ce qui aurait impliqué de recontacter les usagers après quelques semaines voire quelques mois. Par conséquent, nous avons défini l'utilité de la recherche comme: le nombre de documents à consulter (selon une déclaration d'intention de l'utilisateur) sur le nombre de références pertinentes imprimées.

Le projet de recherche

Le projet de recherche fut réalisé en deux étapes:

- 1) une enquête nationale qui avait pour but de décrire la structure et la population des centres de service de télé-référence, et
- 2) une recherche expérimentale ou, disons, « quasi-expérimentale » qui avait pour but de répondre aux trois questions énumérées précédemment et de tester neuf hypothèses de recherche.

Collecte des données

Les données ont été colligées entre le 26 mars et le 4 mai 1979 dans six institutions possédant un centre de

service de télé référence: trois bibliothèques universitaires (Laval, Guelph et Western Ontario) et trois bibliothèques spécialisées (Institut de Recherche de l'Hydro-Québec, Bell Canada et Agriculture Canada). Au total, 34 usagers et 22 médiateurs ont effectué 102 recherches sur quatre systèmes (CAN/OLE, QL SYSTEMS, DIALOG et ORBIT) (voir le tableau 4).

Parmi toutes les variables mesurées, seulement une dizaine furent contrôlées ou, disons, manipulées:

- 1) le type de bibliothèque (universitaire ou spécialisée) afin d'évaluer et d'analyser les différences entre les deux;
- 2) la langue maternelle des usagers et des médiateurs (anglais ou français) afin d'évaluer et d'analyser les différences entre les deux;
- 3) l'expérience préalable des usagers en télé référence (usagers naïfs, usagers occasionnels, usagers experts) afin d'évaluer et d'analyser l'impact de l'expérience sur la performance;
- 4) la formation des usagers (identique pour tous les usagers) afin de neutraliser cette variable;

5) le mode d'interrogation (chaque usager devait effectuer trois recherches: une en mode délégué, une en mode direct et une en mode conjoint) afin d'évaluer et d'analyser les différences entre les trois modes;

6) chaque recherche devait porter sur un sujet différent, afin d'éviter l'effet d'apprentissage;

7) l'ordre d'utilisation des modes d'interrogation pour chaque usager était aléatoire afin, encore une fois, d'éviter un effet d'apprentissage;

8) le laps de temps entre la séance de formation et les recherches devait être identique pour chaque usager;

9) le facteur «coût» était nul, donc constant, pour ne pas créer de pression inutile sur l'usager et, aussi, parce que nous voulions trouver des usagers...; et finalement

10) à cause de problèmes évidents d'apprentissage, le nombre de systèmes utilisés a dû être restreint à quatre.

Les usagers et médiateurs furent choisis par les responsables des centres de service.

Un dossier d'information assez complet était envoyé à chaque participant environ un mois avant les séances de formation et de recherche. Il comprenait aussi trois exemplaires du formulaire de demande de recherche que chaque usager devait compléter avant le projet, un formulaire pour chacune des trois questions de recherche. Les usagers devaient lire ce dossier attentivement *en portant une attention particulière au(x) système(s) qu'ils prévoient utiliser pour leur recherche directe.*

Dans chaque institution, les usagers étaient formés en groupes de trois; le lendemain avait lieu leur journée de recherches.

La journée de formation comprenait une demi-journée théorique pendant laquelle nous expliquons la télé référence et répondons aux questions des usagers; la seconde partie de la journée était réservée aux démonstrations des systèmes et à la pratique de chaque usager sur le (ou les systèmes) qui l'intéressait.

Le lendemain était la journée des recherches. Chaque usager devait apporter ses trois formulaires de demandes d'interrogation.

Ensuite les recherches étaient effectuées selon la procédure normale en

Tableau 4: Coordonnées du projet de recherche

Nom de l'institution	Type de Bibliothèque	Nombre d'usagers	Nombre de recherches	Nombre de médiateurs	DATES (1979)	
					Formation	Recherche
AGRICULTURE CANADA	Spécialisée	5	15	3	26 mars 27 mars	28 mars 29 mars
HYDRO-QUÉBEC	Spécialisée	5	15	1	2 avril 4 avril	3-4 avril 5-6 avril
BELL-CANADA OTTAWA MONTREAL	Spécialisée Spécialisée	3 3	18	1 3	10 avril 7 mai	11 avril 8 mai
UNIVERSITÉ LAVAL	Universitaire	6	18	6	17 avril 19 avril	18 avril 20 avril (un le 19 avril)
UNIVERSITY OF GUELPH	Universitaire	6	18	4	23 avril 25 avril	24 avril 26 avril (un le 30 avril)
UNIVERSITY OF WESTERN ONTARIO	Universitaire	6	18	4	30 avril 2 mai	1 ^{er} mai 3 mai (un le 2 mai)
		34	102	22		

vigueur dans chaque centre de service. Une entrevue préliminaire et subséquente à la recherche était suggérée au besoin (mais non obligatoire) pour les recherches en mode délégué et en mode conjoint. Il n'y avait pas d'entrevue pour les recherches en mode direct.

Il n'y avait aucune restriction imposée quant au nombre de systèmes ou de banques à interroger, ou quant au nombre de références que le chercheur pouvait faire imprimer en direct ou en différé. Le but visé était que les recherches soient le plus naturelles possible et répondent à des besoins réels. Toutefois, tous les usagers, une fois la recherche terminée, devaient indiquer, pour chacune des références imprimées, en direct et en différé, si cette référence était ou non pertinente.

Les données furent colligées au moyen de six questionnaires: le formulaire de demande de recherche, le formulaire de renseignements sur l'utilisateur, le formulaire de renseignements sur le médiateur, le formulaire de données sur la recherche, le formulaire d'évaluation de l'utilisateur et le formulaire d'évaluation du médiateur. En plus, toutes les stratégies de recherche furent imprimées et nous conservions une copie de toutes les références imprimées, en direct et en différé, après leur évaluation par l'utilisateur. Enfin, les factures et relevés statistiques provenant des systèmes furent très utiles dans la compilation des temps et des coûts de recherches.

Résultats

Les résultats de l'étude sont intéressants. Évidemment, de par la définition même des mesures de temps, de coût et de performance — du moins celles qui comprenaient les temps et les coûts totaux de la recherche, il était assez évident que le mode direct allait être avantage par rapport aux deux autres. Toutefois l'un des objectifs de l'étude était précisément de comparer les trois modes d'interrogation pour savoir, notamment, lequel était le moins coûteux, le plus efficace et le plus utile pour une institution donnée, compte tenu que dans la plupart des cas les usagers et les médiateurs étaient des employés de la même institution.

D'autre part, la prudence est de règle dans l'interprétation des résultats. En effet, les usagers utilisés pour l'expérience ne représentent peut-être pas véritablement toute la population des usagers des services de télé référence. Cette population peut être divisée grossièrement en deux catégories, eu

égard au mode d'interrogation: les usagers disposés et disponibles pour participer activement à leurs recherches documentaires et ceux qui ne le sont pas et qui ne s'intéressent qu'au produit fini, c'est-à-dire à l'information demandée. L'enquête préliminaire montrait que les usagers de la première catégorie se retrouvent surtout dans les universités et les bibliothèques gouvernementales tandis que ceux de la seconde se recrutent principalement dans l'industrie et le commerce. C'est pourquoi un échantillon des deux catégories avait été retenu pour l'expérience. Même si très peu de différences réelles furent découvertes aux niveaux des mesures de performance entre les deux types d'institution, il n'en demeure pas moins que ce biais est très réel. Il est en effet évident que les usagers qui ont participé à la présente étude étaient, par définition, intéressés à s'engager activement dans leurs recherches documentaires. Sinon, ils n'auraient pas accepté d'y participer. Par conséquent, la généralisation que nous faisons dans la présente étude sur les usagers, leurs caractéristiques et leur comportement, sont davantage susceptible de s'appliquer aux usagers intéressés à participer activement à leurs recherches documentaires. Les autres seraient plutôt candidats à la délégation de leur recherche.

Comparaison des modes d'interrogation

Ceci dit, voyons maintenant comment les trois modes d'interrogation se sont comportés par rapport aux diverses mesures de performance.

Mode direct

En simplifiant et en schématisant un peu, nous pouvons conclure que le mode direct s'est avéré le meilleur sur la majorité des mesures globales de temps et de coût, c'est-à-dire: l'efficacité du mode d'interrogation, le coût-efficacité du mode d'interrogation, le coût-efficacité du chercheur, l'effort requis dans chaque mode d'interrogation et l'effort requis de l'utilisateur. C'est donc dire que toutes ces mesures furent *minimisées* dans les recherches en mode direct. (voir le tableau 5). Les différences entre les trois modes d'interrogation étaient statistiquement significatives ($P < .05$) pour toutes ces mesures. L'explication la plus plausible de ces résultats est que, pour toutes ces mesures, le mode direct nécessitait moins de temps et donc moins d'argent de la part de l'institution. Le mode direct, en effet ne comprenait aucune entre-

vue préliminaire ou subséquente à la recherche et n'engageait que le temps d'un seul intervenant — l'utilisateur — pendant la recherche. Il était donc tout à fait naturel, si la formation des usagers était adéquate et efficace, que le mode direct révèle les meilleurs résultats pour ces mesures.

D'autre part, au chapitre du coût moyen total d'une recherche, (voir le tableau 6) en tenant compte de tous les coûts — temps de l'utilisateur, temps du médiateur, coûts de connexion, de communication et d'impression —, le mode direct coûtait en moyenne 55,81\$ par recherche, par rapport à 52,51\$ pour le mode délégué et 96,31\$ pour le mode conjoint. C'est donc dire qu'en ne tenant pas compte des coûts de formation, il n'en coûterait pas beaucoup plus cher pour une institution de former les usagers et de les laisser effectuer eux-mêmes leurs recherches au terminal, que de former des médiateurs et de les faire chercher pour ou avec l'utilisateur.

Pour être tout à fait honnête, toutefois, il faut ajouter au coût de la recherche en mode direct celui de la formation des usagers qui sera normalement plus considérable pour l'institution que celui de la formation des médiateurs, beaucoup moins nombreux. Cependant, il nous semble que la formation de chaque usager n'aura pas à être aussi poussée que celle d'un médiateur. D'autre part, quel que soit le coût de formation moyen par usager, il ne reviendra évidemment pas régulièrement pour chacune des recherches effectuées par chaque usager par la suite. Bien au contraire, plus l'utilisateur effectuera de recherches par la suite, plus le coût unitaire de formation s'amortira pour l'institution. Tout dépendra de la fréquence d'utilisation par chaque usager et de la nécessité de le re-former et de le tenir à jour sur les développements et les changements des systèmes.

Une dernière mesure de performance fut optimisée par le mode direct: la satisfaction des usagers. Plus précisément, le mode direct révèle (voir tableau 5) les meilleurs taux globaux de satisfaction des usagers (81,2%) et les meilleurs taux de satisfaction de la pertinence des références (74%) du mode d'interrogation (82,3%) et de l'utilité de leur recherche (87,5%).

Cependant, il faut répéter que la prudence est nécessaire dans l'interprétation de ces résultats. Comme je l'ai expliqué précédemment, il est fort probable que les usagers ayant accepté de participer à l'étude aient

Tableau 5: Mesures de performance: résultats et comparaison des modes d'interrogation.

Mesure	Mode direct (moyennes)	Mode délégué (moyennes)	Mode conjoint (moyennes)	Différence statistiquement significative (P < .05)
Efficacité de la recherche (taux de précision) (SE)	41,4%	42,1%	30,8%	oui
Efficacité de mode d'interrogation (SI)	4,64 min	11,17 min	11,32 min	oui
Efficacité du chercheur (SGE)	2,52 min	2,05 min	2,91 min	oui
Coût-efficacité du mode d'interrogation (CE)	3,99\$	5,90 \$	5,86\$	ou
Coût-efficacité du chercheur (CI)	2,57\$	3,10 \$	3,15\$	oui
Effort requis dans chaque mode d'interrogation (SEF)	1,25\$	2,62 \$	2,88\$	oui
Effort requis de l'utilisateur (USEF)	4,64 min	6,57 min	7,5 min	oui
Satisfaction de l'utilisateur (US):				
nombre de références repérées	71,1%	65,7%	77,0%	oui
nombre de références pertinentes	74,0%	65,6%	69,1%	oui
mode d'interrogation	82,3%	68,6%	80,4%	oui
utilité de la recherche (SU)	87,5%	68,9%	75,8%	oui
télé référence	91,0%	85,7%	92,5%	non
TOTAL	81,2%	70,9%	79,0%	oui
Utilité de la recherche (SU)	75,02%	60,66%	77,44%	oui

été, par définition, davantage portés vers les modes d'interrogation les plus actifs, soit respectivement les modes direct et conjoint. Le taux assez élevé de satisfaction des usagers dans le mode conjoint (79%) confirme d'ailleurs cette hypothèse.

Enfin, le mode direct (voir le tableau 6) est également celui dans lequel le chercheur a utilisé le plus souvent le format d'impression le plus simple (auteur, titre, source); c'est aussi le mode où l'on rapporte les temps de préparation moyens les plus courts et les temps de connexion moyens les plus courts; enfin c'est le mode où les coûts d'impression furent les plus bas. Sans que le phénomène ne soit confirmé ni par l'analyse des questions de recherche, ni par les justifications des usagers pour le choix de chaque mode d'interrogation, ces résultats semblent montrer que les recherches effectuées en mode direct ont peut-être été dans l'ensemble plus simples et moins sophistiquées que dans les deux autres modes. Il n'en demeure pas moins que c'est dans le mode direct que les usa-

gers se sont déclarés le plus satisfaits de leur recherche, de la pertinence des références repérées et surtout, de l'utilité de leur recherche.

Le mode délégué

Comme prévu, le mode délégué s'est avéré supérieur sur les mesures de performance mettant en cause l'expérience des intervenants: le taux de précision (42,1%) et l'efficacité du chercheur (2,05 minutes par référence pertinente) (voir le tableau 5).

Toutefois, c'est également dans ce mode que l'on retrouve le pire taux de coût-efficacité (5,90\$ par référence pertinente imprimée) et les plus bas taux de satisfaction et d'utilité de la recherche. Si l'on peut probablement expliquer le faible taux de satisfaction par le même phénomène qui fait que les modes les plus actifs ont été les plus populaires auprès des sujets de l'expérience, par contre, il est plus difficile d'expliquer les deux autres résultats.

Si l'on considère, par exemple, le taux d'utilité de la recherche, il est très nettement inférieur (60,6%) à celui des

deux autres modes, une seule explication semble possible: la définition même du taux d'utilité (nombre de documents à consulter sur le nombre de références pertinentes imprimées). C'est dans le mode délégué que nous retrouvons, le nombre moyen de références pertinentes imprimées le plus élevé (78,6) (voir tableau 6). Cependant, la proportion des documents signalés qui sera consultée est plus faible même si le nombre total de documents à consulter est plus élevé dans le mode délégué que dans les deux autres. Vraisemblablement donc, l'utilisateur a l'intention de consulter un nombre limité de documents, quel que soit le nombre de documents pertinents qui lui sont signalés. Et ce n'est pas tellement le nombre de références pertinentes qui va satisfaire l'utilisateur (65,6% de satisfaction en mode délégué) ou qui va lui être utile (68,9% de satisfaction en mode délégué) mais bien davantage la pertinence des documents qu'il entend consulter pour satisfaire son besoin d'information documentaire.

Tableau 6: Comparaison des modes d'interrogation sur différentes mesures d'évaluation

Mesure	Mode direct (moyennes)	Mode délégué (moyennes)	Mode conjoint (moyennes)	Différence statistiquement significative (P < .05)
Coût total de la recherche	55,81 \$	52,51 \$	96,32 \$	oui
Nombre de commandes différentes utilisées pendant la recherche	10,8	11,4	12,1	non
Nombre de points d'accès différents utilisés pendant la recherche	2,5	3	3,3	non
Nombre de références examinées	121	154	157	non
Nombre de références imprimées: en direct	19,8	21,7	21,8	non
en différé	52,7	126,1	113,6	oui
au total	72,5	147,8	135,4	oui
Nombre de références pertinentes imprimées	41,8	78,6	45,4	oui
Nombre de documents à consulter	31,5	47,7	35,2	oui
Nombre de banques consultées par recherche	1,85	2,5	2,47	non
Temps de préparation	25,14 min	26,03 min	30,72 min	non
Temps de l'entrevue préliminaire	—	13,24 min	23,6 min	oui
Temps de connexion	37,18 min	27,74 min	47,44 min	oui
Temps de l'entrevue post-recherche	—	3,09 min	1,32 min	oui
Temps total de l'utilisateur	62,32 min	42,36 min	103,08 min	oui
du médiateur	—	44,07 min	72,36 min	oui
Temps total de la recherche	62,32 min	86,43 min	175,44 min	oui
Coûts de l'utilisateur par recherche	14,36 \$	9,24 \$	20,96 \$	oui
Coûts du médiateur par recherche	—	9,29 \$	16,38 \$	oui
Coûts de connexion	32,09 \$	23,53 \$	43,11 \$	oui
Coûts de communications	4,05 \$	3,16 \$	5,28 \$	oui
Coûts d'impression	5,31 \$	7,29 \$	10,59 \$	non
Nombre de recherches pendant lesquelles un thésaurus fut utilisé	3	11	7	oui
Nombre de recherches pendant laquelle fut utilisé le format d'impression:				
le plus simple (auteur-titre-source)	22	15	17	non
complet	16	24	23	non
autre*	16	5	12	oui

*Titre d'abord pour vérifier la pertinence de la référence puis un des deux autres formats ensuite.

Une autre analyse vient confirmer cette hypothèse. En effet, un test de régression multiple sur le taux de satisfaction global des médiateurs révèle que seulement deux variables spécifiques mesurant la satisfaction des médiateurs contribuent à expliquer statistiquement le tiers de leur taux de satisfaction global: il s'agit de leur satisfaction du nombre de références repérées et de leur satisfaction du mode d'interrogation. Ces résultats montrent bien que le médiateur ne possède que très peu d'éléments pour juger seul de la performance d'une recherche. C'est pourquoi bien souvent –et cela est confirmé dans la présente étude– il imprimera davantage de références au cours des recherches en mode délégué et laissera l'utilisateur faire le jugement de pertinence final.

Le mode conjoint

Ce qui nous amène à parler du mode conjoint. Souvent fois qualifié d'heureux compromis entre un mode direct difficile à mettre en pratique et un mode délégué où l'interaction et l'«heuristique» de la recherche sont à toutes fins pratiques perdues, le mode conjoint apparaît plutôt, à la lumière des résultats de la présente étude, comme une situation hybride, très coûteuse et qui ne compense pas entièrement pour l'institution ses coûts élevés par des performances accrues et des résultats meilleurs.

En effet, l'étude montre (voir tableau 5) que le mode conjoint révèle les pires taux moyens de précision (30,8%), d'efficacité du mode d'interrogation (11,32 min par référence pertinente imprimée), d'efficacité du chercheur (2,91 min de connexion par référence pertinente imprimée), de coût-efficacité du chercheur (3,15 \$ de coût de connexion par référence pertinente imprimée), d'effort requis dans chaque mode d'interrogation (2,88 \$ de coût de préparation par référence pertinente imprimée) et d'effort total requis de l'utilisateur (7,5 min par référence pertinente imprimée). Dans l'ensemble donc, le mode conjoint n'optimise aucune des mesures de performance de la recherche. Cependant, il semble optimiser les deux seules mesures de bénéfice de la recherche, soit l'utilité de la recherche (77,44%) et la satisfaction de l'utilisateur (79,0%).

Nous arrivons donc à la constatation pour le moins surprenante que le mode

d'interrogation qui permet à l'utilisateur d'effectuer les recherches les plus satisfaisantes et les plus utiles pour lui est également celui qui coûte le plus cher à son employeur et qui semble le moins efficace.

Nous devons, toutefois, expliquer que la plupart des mesures de performance sont transparentes pour l'utilisateur, c'est-à-dire que ce ne sont pas des mesures importantes pour lui. En général, tout ce que l'utilisateur recherche du service de téléconsultation, c'est de l'information documentaire pertinente à ses besoins. Que la recherche soit utile et satisfaisante lui suffit amplement. Les autres mesures, les mesures de performance, sont utiles à l'institution et à son service de téléconsultation. En ce sens, donc, les résultats de la présente étude sont tout à fait dans l'ordre des choses.

Un dernier point, toutefois, reste à souligner: les recherches effectuées pendant la présente étude étaient gratuites. Le coût du service et sa variation assez marquée d'un mode à l'autre auraient fort probablement été une troisième variable sensible à l'utilisateur. Il serait intéressant en tout cas de vérifier son impact sur le choix d'un mode d'interrogation par l'utilisateur.

En résumé donc, les comparaisons entre les trois modes d'interrogation utilisés dans la présente étude donnent le portrait suivant:

le mode direct optimise la plupart des mesures globales de performance,

le mode conjoint optimise surtout les mesures de bénéfice, et

le mode délégué optimise les seules mesures de performance basées sur l'expérience du chercheur.

Conclusion

En extrapolant un tant soit peu ces observations et en considérant le fait que les résultats obtenus sur les mesures de bénéfice dans le mode direct (81,2% de satisfaction et 75,02% d'utilité) étaient tout de même assez proches de ceux du mode conjoint (79% de satisfaction et 77,44% d'utilité), il est permis de croire qu'avec un minimum de formation de base adéquate, avec des programmes de formation continue faciles d'accès et avec des médiateurs disponibles pour consultation au besoin, les utilisateurs qui le veulent peuvent effectuer directement leurs recherches au terminal avec autant d'efficacité et de profit que les médiateurs.

La seconde partie de cette étude consistait essentiellement à vérifier certaines hypothèses de recherche relatives 1°) aux caractéristiques des questions de recherche (différences entre les modes d'interrogation, corrélation entre les temps de préparation et de connexion, choix d'un mode d'interrogation par l'utilisateur, sélection des systèmes et des banques dans chaque mode d'interrogation, différences dans les temps de préparation, de connexion et de recherche entre les utilisateurs et médiateurs anglais et français et entre les bibliothèques spécialisées et universitaires) 2°) aux caractéristiques des médiateurs (différences dans les mesures de performance entre les médiateurs «humanistes» et les médiateurs «systémistes», différences dans les mesures de performance entre les médiateurs «généralistes» et les médiateurs «spécialistes», corrélations entre l'expérience des médiateurs en téléconsultation et les mesures de performance); 3°) aux caractéristiques des utilisateurs (différences dans les mesures de performance entre les utilisateurs «humanistes» et les utilisateurs «systémistes», différences dans les mesures de performance entre les diverses combinaisons ou interactions possibles des médiateurs et des utilisateurs «humanistes» et «systémistes», choix des caractéristiques des utilisateurs pour l'analyse discriminante).

Enfin, la recherche contenait un troisième volet fort important et relativement novateur: une analyse discriminante ayant pour objectif de tester la possibilité de prévoir, à partir de caractéristiques de l'utilisateur et de sa question de recherche, le meilleur mode d'interrogation à utiliser pour optimiser chacune des mesures de performance.

Les résultats des vérifications d'hypothèse et de l'analyse discriminante ont été présentés au 11^e Congrès annuel de l'Association canadienne des sciences de l'information, à Halifax, les 24-26 mai 1983.

RÉFÉRENCES

- DESCHATELETS, Gilles, *Towards an Optimal Level of Participation of the Human Search Intermediary in the User-System Interface of On-line Bibliographic Search Services* (London, Ontario, School of Library and Information Science, University of Western Ontario, 1982) 703 pages (Ph.D. dissertation).
- DESCHATELETS, Gilles, *Enquête sur les services de téléconsultation au Canada*, (Québec, Université Laval, mars 1980) 142 pages.



LE GRAND DICTIONNAIRE ENCYCLOPÉDIQUE LAROUSSE, voilà l'ouvrage de référence essentiel pour tous ceux qui cherchent des points de repère capables de répondre aux curiosités de plus en plus vastes que suscitent les contradictions du monde contemporain.

- 15 volumes (19 x 28 cm)
- environ 12000 pages
- 180000 articles
- 12000 dessins et schémas
- 11000 photographies
- 1300 cartes
- 60 tryptiques

- avec la collaboration d'environ 1000 spécialistes

- PROFITEZ DE NOS PRIX AVANTAGEUX DE SOUSCRIPTION

LE VOLUME 6 VIENT DE PARAÎTRE

les éditions françaises

DISTRIBUTEURS EXCLUSIFS
POUR LE CANADA
1411, rue Ampère
Boucherville, QC J4B 6C5
Tél.: (514) 641-0514

• TOUTES LES NOUVELLES PARUTIONS, JEUNESSE ET ADULTE, ÉTALÉES PENDANT 4 MOIS.

- **DU NOUVEAU:**
- **SALLE D'EXPOSITION PERMANENTE DE PLUS DE 10 000 TITRES.**
- **UN ÉTALAGE DE JEUX ÉDUCATIFS DE LOGICIELS ET DES DIDACTICIELS.**
- UN SERVICE D'ENVOI DES LISTES DE NOUVELLES PARUTIONS.
- UN ÉVENTAIL COMPLET DE TOUS LES OUVRAGES DE LANGUE FRANÇAISE: QUÉBÉCOIS ET ÉTRANGERS.

Heures d'ouverture: du lundi au mercredi 9h à 18h
jeudi et vendredi 9h à 21h
samedi 9h à 17h

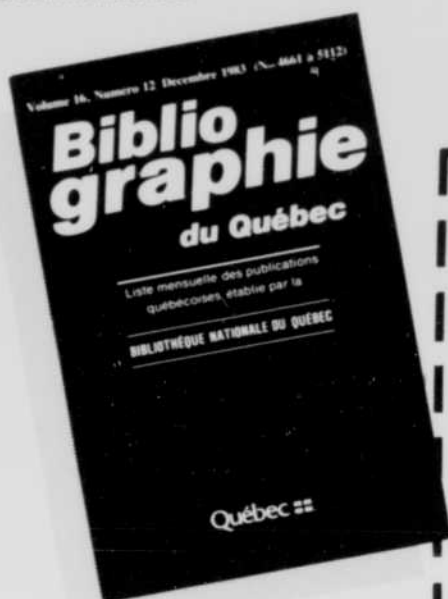
Librairie agréée par le ministère des Affaires culturelles.

**Librairie
FRANCE
QUÉBEC
Inc.**

3550, rue Rachel est
Montréal, Québec
H1W 1A7
Tél.: (514) 526-5951



Gouvernement du Québec
Ministère des
Communications



La **Bibliographie du Québec** regroupe les monographies, les publications en série et les documents cartographiques des secteurs privé et public nouvellement parus et soumis au dépôt légal.

Chaque numéro comprend deux index :

- l'index des auteurs, titres et vedettes secondaires
- l'index des vedettes-matières

L'abonnement d'un an (12 numéros) comprend un index mis à jour au 31 décembre de chaque année.

BON DE COMMANDE (Abonnement)



Quantité	Code EOQ	Titre	Coût annuel	Total
	9970-5	Bibliographie du Québec	75 \$	

À compléter en lettres moulees S.V.P.

Nom _____ Date _____

Adresse _____

Ville _____ Code postal _____

Telephone _____

IMPORTANT

- Les prix sont sujets à changement sans avis préalable.
- Les abonnements entrent en vigueur 15 jours ouvrables après l'encaissement du paiement et sont valables pour une période de 12 mois.
- Pour obtenir les numéros antérieurs à votre abonnement, veuillez communiquer à l'adresse ci-contre.

Retourner ce coupon à :

Ministère des Communications
Direction de la Commercialisation
C.P. 1005
Québec (Québec)
G1K 7B5
Tél. (418) 643-5150

Le numéro **7 \$**

Toute commande est payable à l'avance à l'ordre de **Les publications du Québec**

B757

Québec

INVENTAIRE DES TRAVAUX DE TERMINOLOGIE

La Banque de terminologie du Québec vient de publier la troisième édition d'un document intitulé **Inventaire des travaux de terminologie: projets et travaux en cours, fichiers et publications récentes connus au 1er février 1983.**

Cette réalisation s'inscrit dans le prolongement d'enquêtes précédentes, effectuées en 1974 et 1975. Le succès de ces deux éditions antérieures a incité les responsables à reprendre le travail, afin d'être en mesure de fournir aux personnes intéressées un outil de référence plus actuel, reflétant au maximum les préoccupations de tous ceux que touche la terminologie.

Cette nouvelle édition de l'**Inventaire** présente 430 signalisations, émanant de 121 organismes distincts, représentant l'administration publique québécoise et canadienne, les services de traduction des entreprises québécoises, les organisations nationales et internationales, ainsi que quelques groupes restreints. Il tente de refléter l'ensemble des préoccupations du milieu, même si ses compilateurs sont conscients du fait qu'il est très difficile, sinon impossible, de circonscrire de façon complète tout ce qui peut être produit, ou élaboré dans le cadre de la francophonie.

La nouvelle édition de l'**Inventaire** des travaux de terminologie répond à deux types de préoccupations des terminologues et des traducteurs. En premier lieu, il tente de faire connaître les travaux ou projets de travaux de terminologie à tous les intervenants impliqués dans le domaine. En second lieu, il vise à faciliter le stockage d'une terminologie nouvelle, avant même la publication du document. Nous espérons également que ce travail permettra de mettre en contact des gens œuvrant dans les mêmes sphères d'activités, évitant par le fait même des doublons éventuels.

On peut se procurer des exemplaires de l'**Inventaire des travaux de terminologie** chez l'Éditeur officiel du Québec, au coût de 4,95\$.

L'Office de la langue française compte publier vers la fin de l'année 1984 une nouvelle édition de cet inventaire, et souhaiterait que lui soit signalé tout projet, travail ou fichier de terminologie en préparation, en projet, ou en voie de publication.

Les informations recueillies pourront être acheminées à

Marcel Plourde
Service du traitement des
données terminologiques
Office de la langue française
700, boul. St-Cyrille est
Québec, (Québec)
Canada G1R 5G7

PROTOCOLE DE RÉDACTION

Argus vise à assurer l'information et le développement professionnel des membres de la Corporation des bibliothécaires professionnels du Québec. Les articles publiés traitent de la formation, du rôle et du statut du bibliothécaire dans la société, du professionnalisme, des nouvelles orientations et de la recherche en bibliothéconomie et en science de l'information, de l'apport des autres disciplines.

Le comité de rédaction accepte des articles originaux soit en français, soit en anglais. À l'occasion il publie des textes de conférences prononcées dans le cadre des activités de la Corporation.

Les publications peuvent prendre plusieurs formes:

- articles de fond (de 15 à 25 pages dactylographiées)
- textes plus courts rendant compte d'un événement récent
- entrevues
- chroniques sur l'audiovisuel, les revues professionnelles, la recherche (de 3 à 6 pages dactylographiées)
- lettres à l'éditeur paraissant sous le titre "Argus commentaires" et commentant une question d'actualité ou un article récemment paru dans la revue
- comptes rendus d'ouvrages sur le professionnalisme ou la recherche.

Les articles doivent être soumis en quatre exemplaires, dactylographiés à interligne double sur papier 21 x 28cm et accompagnés d'un résumé informatif.

Sur la première page doivent figurer le nom, le titre académique, le statut professionnel et le lieu de travail de l'auteur ainsi que le titre et le résumé du texte.

Les notes infrapaginales doivent être numérotées et tapées sur une feuille à part. Il appartient aux auteurs de fournir les références bibliographiques complètes et présentées de la façon suivante:

Monographie:

Herbert Goldhor, *Introduction to Scientific Research in Librarianship* (Urbana, Ill., University of Illinois, 1972), p.198.

Chapitre d'un livre:

William M. Potter, "History, the Behavioral Studies and the Science of Man" in Mary Lee Bundy and Paul Wasserman, *Reader in Research Methods for Librarianship* (Washington, Microcard Editions, 1970), p. 36.

Article de périodique:

Yves Courrier, "Analyse et langage documentaires", *Documentaliste*, vol. 13, no 5-6 (septembre-décembre 1976), pp. 178-189.

Les correcteurs se réservent le droit de renvoyer aux auteurs les textes dont les références ne correspondent pas à ces exigences. Ils peuvent effectuer des corrections mineures des textes sans en avertir les auteurs.

Les auteurs sont priés de conserver un double de leur article. Aucun manuscrit ne leur sera remis.

Chaque auteur recevra trois exemplaires du numéro auquel il aura contribué.

EDITORIAL STATEMENT

Argus aims to inform members of the Corporation of Professional Librarians of Québec and further their professional development. Articles cover the training of librarians and their role and status in society, professionalism, new directions and research in library and information science, and contributions from other fields.

The Editorial Board accepts original articles written in French or in English. Papers delivered at Corporation related functions are published from time to time.

Published items may take several forms:

- in-depth articles (15 to 25 typed pages)
- shorter articles covering a recent event
- interviews
- columns on audio-visual material, professional journals, research (3 to 6 typed pages)
- letters to the Editor, under the heading "Argus commentaires", which discuss a topical issue or an article recently published in the journal
- reviews of works about the profession or research

Articles should be submitted in quadruplicate and be typed (double-spaced) on sheets 21 by 28 cm. An informative résumé should be provided.

The author's name, his/her professional status and place of work, as well as the title and résumé of the article, should appear on the first page.

Foot-notes should be numbered and typed on a separate sheet. The author must provide complete bibliographic references in accordance with the following examples:

Monograph:

Herbert Goldhor, *Introduction to Scientific Research in Librarianship* (Urbana, Ill., University of Illinois, 1972), p. 198.

Chapter of a book:

William M. Potter, "History, the Behavioral Studies and the Science of Man" in Mary Lee Bundy and Paul Wasserman, *Reader in Research Methods for Librarianship* (Washington, Microcard Editions, 1970), p. 34.

Article in a periodical:

Yves Courrier, "Analyse et langage documentaires", *Documentaliste*, vol. 13, no 5-6 (septembre-décembre 1976), pp. 178-189.

Editors reserve the right to return articles if references do not conform to these standards. Minor corrections may be made to an article without the author's prior knowledge.

We request that the author keeps a copy of his/her article. Manuscripts will not be returned.

Authors receive 3 copies of each issue in which their article appears.

