

La numérisation des documents administratifs

Méthodes et recommandations

Septembre 2009

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2010

ISBN : 978-2-550-59634-9

Table des matières

Table des matières.....	2
Introduction.....	4
1.1 Inscription au calendrier de conservation.....	6
1.2 Intégrité et valeur juridique.....	7
1.3 Documentation du processus de numérisation.....	8
1.4 Destruction et conservation des documents sources.....	9
1.5 Respect du droit d'auteur.....	9
1.6 Accès à l'information.....	10
1.7 Protection des renseignements personnels et confidentiels.....	10
2. Projet de numérisation.....	12
2.1 Définition de projet : les objectifs recherchés.....	12
2.2 Choix des documents à numériser et échéancier du projet.....	12
2.3 Analyse des besoins et des ressources disponibles.....	13
2.4 Numérisation des documents.....	14
2.4.1 Préparation matérielle des documents.....	14
2.4.2 Réalisation de projets pilotes.....	15
2.4.3 Numérisation des documents.....	15
2.4.4 Contrôle de la quantité et de la qualité.....	15
3. Exigences archivistiques.....	18
Introduction.....	18
3.1 Calendrier de conservation.....	18
3.1.1 Mise à jour.....	18
3.1.2 Gestion de la documentation et des déclarations de numérisation.....	18
3.2 Organisation, identification et classification des fichiers.....	19
3.2.1 Classification et classement.....	19
3.2.2 Identification, repérage et métadonnées.....	20
3.3 Conservation et accès aux documents.....	20
3.4 Destruction sécuritaire des documents papier.....	21
3.4.1 Méthodes de destruction sécuritaire.....	21
3.4.2 Pratiques sécuritaires.....	22
4. Exigences techniques.....	23
4.1 Les outils de numérisation.....	23
4.2 Description des outils de numérisation.....	23
4.2.1 Ordinateur.....	23
4.2.2 Écran à haute résolution et carte vidéo.....	24
4.2.3 Numériseur.....	25
4.2.4 Logiciel de numérisation.....	26
4.2.5 Logiciel de reconnaissance optique des caractères.....	27
4.3 Résolution et mode de l'image.....	27
4.4 Les formats de conservation des fichiers.....	28
4.4.1 TIFF non compressé.....	28
4.4.2 PDF/A-1.....	28
4.4.3 XML.....	29

4.5 Les supports de stockage et la protection des fichiers	30
4.5.1 Supports optiques	30
4.5.2 Supports magnétiques	31
4.5.3 Choix de formats et de supports de conservation.....	31
5. Aide-mémoire pour la numérisation de documents	32
5.1 Cadre juridique - obligations légales et recommandations	32
5.2 Projet de numérisation - recommandations.....	33
5.3 Exigences archivistiques - recommandations	33
5.4 Exigences techniques - recommandations	33
Bibliographie.....	35
Annexe 1 - Déclaration de numérisation d'archives.....	37

Introduction

Dans un monde où de plus en plus de documents sont produits ou diffusés en mode numérique, la numérisation de documents sur support analogique, notamment sur papier, est en croissance dans les organisations depuis plus de 25 ans. Les améliorations techniques des équipements et la réduction de leurs coûts d'acquisition au cours des dernières années ont contribué à populariser le transfert d'information analogique sous forme numérique. Longtemps utilisée principalement à des fins de diffusion, la numérisation gagne en popularité comme élément d'un programme de gestion documentaire.

Ce petit guide se veut un outil pratique qui n'a nullement la prétention de faire le tour complet de la question de la numérisation des documents. Il s'inscrit dans le prolongement du rôle conseil dévolu à la Direction générale des archives de Bibliothèque et Archives nationales du Québec (BAnQ) auprès des ministères et organismes gouvernementaux (M/O) du Québec. Il vise à fournir aux M/O des balises pour réaliser des projets de numérisation dans un cadre de gestion documentaire conforme à l'actuelle réalité juridique québécoise. Il remplace le *Guide d'imagerie numérique* publié par les Archives nationales du Québec en 1998 et le questionnaire Numérisation de *documents inscrits au calendrier de conservation*.

Le guide aborde l'activité de numérisation sous l'angle de la gestion de projet et de la gestion documentaire. Dans cette optique, il se veut un outil pouvant favoriser la conservation numérique des documents et, dans certains cas et sous certaines conditions, l'élimination des documents source sur supports papier ou microfilm. Il comprend en outre un tableau et des exemples de déclaration de numérisation qui peuvent être utilisés par les M/O pour se conformer aux recommandations de BAnQ et à la législation en vigueur. Le tableau peut également servir de liste de validation lors de la réalisation d'un projet de numérisation. Les exemples de déclarations tiennent compte des principes énoncés dans la *Loi concernant le cadre juridique des technologies de l'information*.

Le guide compte cinq sections qui couvrent les divers volets d'une opération de numérisation dans un cadre de gestion documentaire. On y aborde le cadre juridique qui oriente une opération de numérisation au Québec. Une attention particulière est portée à l'impact de la *Loi concernant le cadre juridique des technologies de l'information*, qui établit les principes d'intégrité et d'équivalence fonctionnelle des documents.

La numérisation est ensuite présentée comme un projet comprenant des étapes bien déterminées et qui doit respecter des exigences archivistiques spécifiques pour être un outil performant dans le cadre d'un programme de gestion intégrée des documents. Par la suite, les exigences et les normes

techniques à prendre en compte pour réaliser un projet de numérisation sont exposées.

Finalement, la dernière section rappelle les obligations légales ainsi que les principales recommandations de BAnQ applicables aux projets de numérisation et de destruction des documents source analogiques.

Le terme *document numérisé*, utilisé tout au long de ce guide, fait référence au document numérique qui résulte de la numérisation d'un document analogique, notamment sur support papier.

Ce guide est destiné en priorité au personnel des organismes publics assujettis à la *Loi sur les archives*. Nous espérons qu'il pourra aussi rendre service aux autres personnes intéressées ou concernées par la numérisation des documents administratifs.

1. Cadre juridique

Pour mettre en œuvre une activité de numérisation, il faut connaître les obligations légales qui s'y appliquent et les respecter. Dans cette partie, on trouvera les principales composantes du cadre juridique dont il faut tenir compte.

1.1 Inscription au calendrier de conservation¹

Numérisation des documents entraîne la **révision des règles de conservation** des séries documentaires touchées par cette opération. En effet, il est nécessaire :

- de réviser la règle de conservation du document source pour inscrire le nouveau support au calendrier de conservation et revoir les périodes d'utilisation à l'état actif et semi-actif, ainsi que le mode de disposition finale des documents;
- Créer une règle de conservation spécifique pour la documentation de numérisation ou ajouter cette dernière parmi les types de documents de la série documentaire faisant l'objet de la numérisation.

Les organismes publics doivent obligatoirement soumettre ces règles de conservation à Bibliothèque et Archives nationales du Québec (BAnQ) aux fins d'approbation, et ce, avant la destruction des documents source et, préférablement, avant le début du projet de numérisation. C'est uniquement par cette approbation qu'il devient possible de détruire les documents sources visés par une opération de numérisation.

L'exemple ci-dessous illustre la manière dont doit être inscrit au calendrier de conservation un changement de support lié à une numérisation. Il faut s'assurer que la durée active des documents sources inclue la période requise à la validation exhaustive du contrôle de la qualité et doit être indiqué dans la règle.

Avant le transfert de support :

DÉLAI DE CONSERVATION						
Numérotation	Exemplaire	Supports de conservation		Période d'utilisation des documents		Disposition
				Actif	Semi-actif	Inactif
	Principal	PA		2	3	C
	Secondaire	PA		2	0	D
REMARQUES RELATIVES AU DÉLAI DE CONSERVATION						

¹Loi sur les archives (L.R.Q., c. A-21.1, art. 7, 8).

Transfert de support

DELAI DE CONSERVATION								
Numérotation	Exemplaire	Supports de conservation		Période d'utilisation des documents			Disposition	
				Actif	Semi-actif	Inactif		
	Principal	PA		888	R1	0	D	
	Principal	DM		2		3	C	
	Secondaire	PA, DM		2		0	D	

REMARQUES RELATIVES AU DÉLAI DE CONSERVATION
R1 : Les documents sont numérisés et détruits après la validation exhaustive du contrôle de la qualité.

1.2 Intégrité et valeur juridique²

Un document numérisé dont l'intégrité est assurée possède une **valeur juridique**, au même titre que le document source analogique, s'il répond à l'une des conditions suivantes :

- il est susceptible de produire des effets juridiques (ex. : un contrat);
- il est susceptible d'être admis en preuve devant un tribunal ou un arbitre pour faire valoir un droit.

L'**intégrité** d'un document, élément qui constitue le fondement de sa valeur juridique, est assurée lorsqu'on peut vérifier que :

- l'information est intégrale et n'a pas été altérée;
- le support de l'information lui assure stabilité et pérennité.

Il faut donc maintenir l'intégrité du document numérisé pendant tout son cycle de vie en mettant en place des mesures de sécurité pour le protéger.

La notion d'intégrité, telle que définie par la *Loi concernant le cadre juridique des technologies de l'information (LCCJTI)*, est aussi associée à la nécessité d'appliquer des mesures de sécurité lors du processus de transfert. De plus, afin de pouvoir facilement démontrer l'intégrité du document numérisé, il faut s'assurer que le support utilisé pour la conservation offre une garantie de stabilité et de pérennité maximale suivant les normes en vigueur.

Deux exemplaires d'un même document sur des supports différents jouissent de **l'équivalence fonctionnelle**, c'est-à-dire qu'ils peuvent remplir les mêmes fonctions. Le choix d'un support ou d'une technologie pour consigner de

²Loi concernant le cadre juridique des technologies de l'information (L.R.Q., c. C-21.1, art. 5, 6, 9, 10, 12, 14, 68).

l'information est libre. Le support est ainsi une caractéristique neutre par rapport à la valeur juridique du document.

Comme l'indique la loi, le document numérisé peut remplacer le document source sur support analogique. De cette façon, on peut utiliser aux mêmes fins le document numérisé et le document analogique et, si nécessaire, un document peut servir à reconstituer l'autre. Dans la mesure où son intégrité est assurée, le document numérisé peut remplir les fonctions d'un original³.

La *LCCJT* précise que, lors d'un transfert d'information d'un support vers un autre, par exemple lors d'une numérisation, l'intégrité du document est maintenue, même en présence de différences touchant :

- l'emmagasinage de l'information;
- la présentation de l'information;
- l'information sur le support ou la sécurité;
- la pagination;
- le caractère tangible ou intangible des pages;
- le format des pages;
- la présentation recto ou verso;
- l'accessibilité des pages en tout ou en partie;
- le repérage séquentiel ou thématique de l'information.

1.3 Documentation du processus de numérisation⁴

Pour assurer l'intégrité et la valeur juridique d'un document résultant d'un transfert de support, il est obligatoire de documenter le processus de numérisation.

La **documentation exigée** doit comprendre, **au minimum** :

- le format du document source;
- le procédé de transfert utilisé;
- les garanties offertes par le procédé de numérisation quant à la préservation de l'intégrité des documents.

Cette documentation doit être **conservée durant tout le cycle de vie du document** numérisé. Elle peut être jointe au document résultant du transfert, à ses éléments structurants ou à son support.

3. *Loi concernant le cadre juridique des technologies de l'information* (L.R.Q., c. C-1.1, art. 12).

4. *Loi concernant le cadre juridique des technologies de l'information* (L.R.Q., c. C-1.1, art.17).

1.4 Destruction et conservation des documents sources⁵

Après une opération de numérisation, deux actions sont possibles : détruire ou conserver des documents source. En général, les documents qui ont fait l'objet d'un transfert de support peuvent être détruits lorsque l'intégrité de l'information est assurée et que quelques autres considérations sont respectées.

Avant de procéder à la destruction, il faut, outre contrôler l'intégrité, avoir documenté et validé l'opération de transfert, respecté le calendrier de conservation et, pendant l'opération, avoir protégé les renseignements personnels ou confidentiels. La destruction des documents contenant des renseignements personnels ou confidentiels doit être faite de façon sécuritaire, en appliquant les mesures décrites dans la section 3.4- *Destruction sécuritaire des documents papier* du présent guide.

Toutefois, certains documents source doivent être conservés sur leur support d'origine. Il s'agit des documents qui, sur leur support d'origine, possèdent une valeur archivistique, historique ou patrimoniale suivant des critères déterminés par règlement du gouvernement, mais qui n'ont pas encore été développés ni publiés. Il convient donc d'être prudent avant de détruire des documents dont le support d'origine peut présenter un intérêt du point de vue de la conservation.

1.5 Respect du droit d'auteur⁶

La numérisation consiste à reproduire un document en le convertissant d'un format analogique à un format numérique. La reproduction étant, au même titre que la communication au public, un droit exclusif de l'auteur, il faut s'assurer, avant de procéder à la numérisation, de détenir les droits d'auteur des documents à reproduire ou d'avoir recueilli le consentement du titulaire des droits. Ces dispositions s'appliquent à tous les documents de tous les organismes.

Dans le contexte de la gestion des documents administratifs, il est **d'usage courant** de reproduire les documents administratifs sans demander l'autorisation du détenteur du droit d'auteur quand :

- ils ont été transmis à l'organisme, sans restriction de reproduction, dans le cadre de l'exercice de ses mandats et responsabilités;
- ils sont reproduits et seront utilisés exclusivement à des fins internes de traitement ou de prise de décision.

⁵Loi sur les archives (L.R.Q., c. A-21.1, art. 18).

⁶Loi sur le droit d'auteur (L.R.Q., c. C-42 art. 3).

1.6 Accès à l'information⁷

Les documents des organismes publics, quel que soit leur support, doivent être identifiables et repérables de façon à pouvoir être communiqués en tout temps. Les documents numérisés n'échappent pas à cette obligation. Ils doivent, comme tout autre document, être classés selon le plan de classification de l'organisme public pour être repérables et rapidement communiqués.

1.7 Protection des renseignements personnels et confidentiels⁸

Les documents à numériser peuvent contenir des renseignements confidentiels, en particulier des **renseignements personnels** qui doivent être protégés et qui ne peuvent être communiqués sans autorisation, par exemple :

- le nom d'une personne physique associé à une information permettant de l'identifier;
- les documents d'un dossier médical;
- les documents d'un dossier d'employé.

Plusieurs articles de loi précisent ce que sont, ou ne sont pas, les renseignements personnels à protéger : en général, le nom d'une personne physique n'est pas considéré comme un renseignement personnel s'il n'est pas accompagné d'une autre information. Par ailleurs, certains renseignements personnels concernant un membre d'un organisme public ont un caractère public.

En plus des renseignements personnels, d'autres renseignements détenus par les organismes publics peuvent être traités de manière confidentielle. Il s'agit des renseignements ayant des incidences sur :

- les relations intergouvernementales;
- les négociations entre organismes publics;
- l'économie;
- l'administration de la justice et la sécurité publique;
- les décisions administratives ou politiques;
- la vérification.

Afin de limiter l'accès aux renseignements personnels et confidentiels, il est nécessaire, avant la numérisation, d'adopter des règles d'habilitation. De

⁷Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels (L.R.Q., c. A-2.1, art. 1, 9, 16, 16.1).

⁸Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels (L.R.Q., c. A-2.1, art. 18 à 41, 54, 55, 56, 57, 59, 63.1).

même, il faut instaurer, pendant et après la numérisation, des mécanismes de contrôle physiques d'accès aux documents et aux installations de numérisation.

Enfin, il faut appliquer des méthodes de destruction sécuritaires aux documents contenant des renseignements personnels. Pour obtenir de plus amples renseignements sur les méthodes de destruction sécuritaires, veuillez consulter la partie du présent guide intitulée *3.4- Destruction sécuritaire des documents papier*.

2. Projet de numérisation

Pour réussir un projet de numérisation, il faut suivre une série d'étapes allant de la conception à la réalisation, en passant par l'analyse des besoins et le respect du cadre législatif. On trouvera ici la description de ces principales étapes.

2.1 Définition de projet : les objectifs recherchés

La première étape de tout projet de numérisation consiste à en définir les objectifs et les buts. On doit se poser les questions suivantes : pourquoi veut-on numériser des documents? Pour un projet de diffusion? Pour protéger et conserver de l'information? Pour économiser de l'espace d'entreposage?

Répondre à ces questions est très important, car c'est ce qui va déterminer non seulement le type de projet à effectuer ainsi que le choix du format et de la résolution des documents mais aussi les coûts associés au projet, sa durée et les ressources nécessaires pour le réaliser. En effet, si on numérise un document en vue d'une conservation permanente, il doit avoir une qualité supérieure (taille, résolution, format) aux documents numérisés à de simples fins de diffusion, sur Internet par exemple. Ceci entraîne forcément des frais supplémentaires. Au besoin, on peut facilement produire, pour diffusion, une version de qualité moindre à partir d'un fichier numérique de conservation, mais l'inverse n'est généralement pas possible. Il est vrai que numériser un document avec une résolution supérieure à ce qu'exige la diffusion, et dans un format qui assure sa conservation permanente, entraîne des frais supplémentaires mais, à moyen et long terme, on économise puisqu'on évite ainsi de numériser plusieurs fois le même document.

Prendre les décisions appropriées en cette matière exige une bonne connaissance du potentiel d'utilisation des documents ainsi qu'un plan de numérisation détaillé, et soutenu par la direction de l'organisation. Bien définir un projet de numérisation, c'est non seulement savoir où l'on va, c'est aussi l'intégrer adéquatement aux priorités institutionnelles.

2.2 Choix des documents à numériser et échéancier du projet

Après avoir cerné les objectifs du projet, il faut déterminer quels documents, ou séries documentaires, seront numérisés, dans quel ordre et combien de temps durera le projet. Il faut situer chaque projet de numérisation dans une perspective à long terme, c'est-à-dire prévoir les types et la quantité de documents à numériser au cours des mois et des années à venir, déterminer ceux qui seront numérisés simplement à des fins de diffusion et ceux qui le

seront à des fins de conservation. Une telle planification permet d'établir des priorités à court et à long terme et de prévoir les besoins qui y sont associés.

Lorsque la numérisation répond à des impératifs de gestion documentaire (documents actifs ou semi-actifs) ou de conservation à long terme (documents inactifs), le choix des documents à numériser doit se faire prioritairement en fonction des délais prescrits par le calendrier de conservation. Il faut donner la priorité aux séries documentaires volumineuses qui doivent être conservées un certain temps au semi-actif ou de manière permanente, plutôt qu'à d'autres séries qui, selon le calendrier de conservation, doivent être éliminées dans un avenir rapproché. La fréquence d'utilisation, l'état de dégradation, l'obligation de diffusion ainsi que les restrictions à l'accès de certains documents peuvent également servir de critères de choix.

Il n'est généralement pas pertinent de numériser toutes les séries documentaires d'une organisation. Il est plutôt recommandé d'assurer d'abord la conservation à long terme des documents nés numériques, puis de numériser les documents papier seulement lorsque les gains en temps, en espace et en argent sont importants, du point de vue des activités, de la conservation ou de la diffusion.

2.3 Analyse des besoins et des ressources disponibles

L'étape suivante consiste à analyser les besoins liés à la réalisation du projet ainsi qu'à dresser l'inventaire des ressources matérielles, humaines et financières disponibles à l'interne. Lors de l'analyse, il faut prendre en compte les compétences particulières du personnel, les équipements déjà acquis, les espaces de travail disponibles, les autres projets en cours au sein de l'organisme, etc. En outre, il ne faut pas sous-estimer les besoins en formation et les coûts récurrents de maintenance et de mise à jour des équipements.

Il faut également porter une attention particulière aux normes et formats de numérisation, ainsi qu'aux supports sur lesquels seront conservés les documents numérisés. Pour les documents destinés à une conservation permanente, le recours à des formats ouverts et non compressés, à une résolution élevée, ainsi qu'à des supports reconnus pour la conservation à long terme peuvent influencer le choix de l'équipement et des logiciels nécessaires à l'opération de numérisation et, par conséquent, avoir des effets sur le budget. Pour plus de renseignements sur les formats et supports de conservation, veuillez consulter les sections 4.5 et 4.6 du présent guide.

C'est en prenant en considération l'ensemble des éléments mentionnés ci-dessus (objectifs, types de documents, délais de conservation, coûts, etc.) qu'il devient possible de déterminer l'ampleur, le type, le coût et la durée d'un projet de numérisation. Cet examen permet aussi de juger de la pertinence d'acquérir ou de louer des équipements pour réaliser la numérisation à l'interne, ou

encore d'en confier la réalisation à un fournisseur externe. Une approche mixte (interne/externe) est également possible.

2.4 Numérisation des documents

Après avoir cerné les priorités et les besoins et avoir pris les dispositions nécessaires pour y répondre, on passe à l'étape de la numérisation proprement dite des documents. Elle comporte un ensemble de sous-étapes et d'actions, communes à tout projet de numérisation ou particulières aux documents d'archives. Les éléments communs à tous les projets sont décrits ci-dessous; ceux qui relèvent de considérations et pratiques archivistiques sont présentés dans la section 3- *Exigences archivistiques*.

L'opération de numérisation peut être effectuée à l'interne ou déléguée à un fournisseur externe. Il n'est pas rare de recourir simultanément à l'une et l'autre formule pour répondre à l'ensemble des besoins. Les principaux avantages de la numérisation à l'interne sont : une meilleure accessibilité aux documents durant le projet, la réduction des risques liés à la manipulation des documents fragiles, un meilleur contrôle de la qualité et le développement d'une expertise institutionnelle. Le recours à l'expertise d'une firme présente lui aussi des avantages, comme l'accès à du personnel possédant une plus grande expérience et à du matériel spécialisé, la rapidité d'exécution, un coût fixe par page numérisée, des prix établis en fonction du volume et l'absence de frais liés à l'obsolescence technologique et à la formation du personnel. Recourir à l'une ou l'autre de ces approches dépend des impératifs de chaque projet et des ressources disponibles.

Que la numérisation soit effectuée à l'interne ou à l'externe, il faut respecter les étapes suivantes :

- Préparation matérielle des documents avant et après la numérisation;
- Réalisation de projets pilotes et apport de correctifs, au besoin ;
- Numérisation des documents ;
- Vérification de la qualité en cours de réalisation et après.

2.4.1 Préparation matérielle des documents

Avant toute chose, il faut préparer les dossiers et les documents à numériser pour faciliter l'opération de numérisation et réduire les risques de détérioration des documents et des équipements. Il faut classer les pages dans le bon ordre, déplier les documents, retirer les reliures à anneaux, éliminer les doubles, retirer les agrafes, trombones, *post-it*, etc. C'est également à ce moment que, au besoin, seront comptées les pages à numériser.

Après la numérisation, il faut prendre le temps de réorganiser les documents, soit pour permettre leur réutilisation, soit pour pouvoir retracer rapidement un document qui doit être numérisé de nouveau pour une raison ou une autre.

2.4.2 Réalisation de projets pilotes

Avant d'entreprendre la numérisation, il est important de réaliser des projets pilotes pour tester les équipements, les méthodes de contrôle des résultats, les mécanismes de normalisation, d'ajout de métadonnées et de création de copies de sécurité des données, le rythme de réalisation, etc. Lors de ces essais, il est recommandé de ne pas utiliser des documents originaux pouvant être altérés par de multiples expositions à la lumière et des manipulations répétées.

En fonction des résultats des essais, il faut apporter les correctifs nécessaires et effectuer ensuite de nouveaux essais jusqu'à ce que l'on ait atteint le niveau de qualité recherché. Bien que cette étape puisse parfois paraître longue et fastidieuse, elle s'impose pour assurer la réussite du projet.

2.4.3 Numérisation des documents

La numérisation doit être réalisée en s'assurant que l'on respecte les règles d'accès et de sécurité et que l'on manipule les documents source de manière à ne pas les endommager.

2.4.4 Contrôle de la quantité et de la qualité

Au cours des différentes étapes de la numérisation, et avant de détruire les documents source, il est essentiel d'évaluer les résultats. Le contrôle comporte deux volets : le contrôle de la quantité et celui de la qualité.

Contrôle de la quantité des documents numérisés

Il est impératif de s'assurer que chacune des pages a été numérisée. Pour ce faire, il est possible de recourir à des techniques manuelles ou automatisées. La méthode traditionnelle consiste à compter manuellement les pages avant la numérisation, puis à comparer le résultat obtenu avec le nombre de pages effectivement numérisées. En ce qui concerne les méthodes automatiques, certains numériseurs et logiciels de numérisation offrent des possibilités de marquage et de comptage des pages. Ces procédés sont plus rapides que le comptage manuel, mais ne dispensent pas de la nécessité de vérifier que toutes les pages ont bien été numérisées. Le marquage des pages n'est pas appropriée pour les documents recto-verso ni pour les documents dont on souhaite conserver l'original intact.

Contrôle de la qualité de documents numérisés

Lors de la numérisation de documents textuels, il est essentiel de contrôler l'intégrité du contenu du document. Il faut s'assurer que le texte (et les annotations, le cas échéant), est bien lisible, qu'aucune phrase ou mot n'est coupé, que la résolution est suffisante, etc. En bref, il s'agit de s'assurer que la reproduction numérique est fidèle au document source et que sa résolution est suffisante en fonction du type de document et de l'usage qu'on prévoit en faire.

En plus de contrôler la qualité du contenu des documents numérisés, il faut contrôler la qualité de la reproduction (luminosité, contraste, couleurs, etc.). L'utilisation de points de repère, grâce à un calibrage et à un étalonnage de l'équipement, tout comme la visualisation d'images en galerie avec des informations contextuelles, permettent de plus facilement valider les résultats et respecter les normes convenues. De plus, cela permet de déceler des erreurs récurrentes telles que :

- la perte de détails dans les points d'ombres et les éléments éclairés;
- l'apparition d'artéfacts d'acquisition ou de compression (bruit, pixellisation);
- une mauvaise luminosité et un mauvais contraste;
- des variations dans l'équilibre des teintes;
- la sous-exposition ou la surexposition des images;
- des images floues, mal orientées ou mal découpées;
- des erreurs de mode (ex. : photographie couleur transformée en noir et blanc).

Dans les cas de travaux de numérisation associés à une substitution de support à des fins administratives, judiciaires ou de conservation, une vérification visuelle exhaustive de chacun des documents numérisés doit être effectuée pour documenter et assurer l'intégrité des reproductions.

Lorsque la numérisation est réalisée à l'externe, le détenteur des documents doit s'assurer que le fournisseur respecte les normes de qualité et de contrôle convenues, qui doivent avoir été indiquées dans le cahier des charges et le contrat d'exécution.

Numérisation et reconnaissance optique de caractères

Sauf pour les manuscrits (dans l'état actuel de la technologie), on peut appliquer à toute image numérisée d'un document textuel une reconnaissance optique de caractères (OCR). Cette reconnaissance peut être réalisée lors de la numérisation, ou après. Par reconnaissance des formes, on entend l'opération par laquelle un logiciel de lecture compare les images des

caractères présents dans le texte avec celles de ses dictionnaires. Il associe ensuite les formes aux caractères appropriés. Le résultat peut être sauvegardé dans un fichier en mode texte (.txt, .doc, .rtf, etc.), ce qui permet d'indexer le contenu du document, d'y effectuer des recherches ou de le modifier.

L'OCR peut s'avérer très utile, surtout pour l'accès aux textes, mais son emploi pour la constitution d'exemplaires de conservation numériques à partir d'un document papier doit être fait avec précaution, puisque l'intégrité du contenu du document doit être assurée. Le cas échéant, en plus des vérifications signalées ci-dessus, il faut effectuer une relecture attentive des fichiers texte générés par le logiciel d'OCR puisque le procédé de reconnaissance de formes n'est évidemment pas parfait. Il est important de s'assurer que le logiciel a bien reconnu tous les caractères, que les accents et les éléments de ponctuation sont bien présents, etc. Enfin, il faut sauvegarder les documents dans un format reconnu pour la conservation permanente (voir section 4.5).

Il est à noter qu'il est possible de conserver le fichier image comme fichier principal de conservation et de n'utiliser le résultat de la reconnaissance optique des caractères qu'à des fins d'indexation et de recherche. Si des résultats de recherches fiables sont attendus, une vérification exhaustive de la reconnaissance doit être effectuée. Réaliser la lecture optique des textes et assurer un contrôle serré des résultats peut entraîner des frais considérables, allant jusqu'à multiplier par dix le coût d'un projet de numérisation. Il est donc important de bien évaluer les besoins d'une telle opération.

3. Exigences archivistiques

Introduction

La réussite d'un projet de numérisation repose non seulement sur une bonne analyse des besoins et sur de bons choix technologiques mais également sur sa conformité aux principes et aux exigences de l'archivistique.

3.1 Calendrier de conservation

3.1.1 Mise à jour

La création de documents numériques dans le cadre de projets de numérisation nécessite l'inscription de nouveaux supports de documents pour la ou les règles du calendrier de conservation. Les modifications du calendrier de conservation peuvent, par exemple, modifier la durée de vie des documents analogiques, ajouter de nouveaux supports, créer de nouvelles règles touchant la documentation du processus de numérisation, etc.

Les modifications apportées au calendrier de conservation doivent être obligatoirement approuvées par BAnQ avant la destruction des documents source mais, préférablement, avant le début du projet de numérisation.

Pour obtenir de plus amples renseignements sur la soumission des calendriers de conservation à BAnQ, et sur leur approbation, veuillez consulter la section : *Accueil / Services / Services spécialisés en archivistique et gestion documentaire / Services aux organismes publics* sur le portail de BAnQ, à l'adresse suivante : <http://www.banq.qc.ca/portail/dt/services/archivistique/ged/public/public.jsp>

3.1.2 Gestion de la documentation et des déclarations de numérisation

Tel qu'indiqué à la section 1.3, la documentation du processus de numérisation doit être conservée aussi longtemps que les documents qu'elle vise.

L'inscription des documents numérisés au calendrier de conservation peut se faire de deux manières, soit par la création d'une règle spécifique pour cette série documentaire, soit par leur ajout comme types de documents aux règles des documents originaux. Dans les deux cas, il importe que le lien entre les documents qui ont été numérisés, la documentation du processus de numérisation et les déclarations soit maintenu tout au long du cycle de vie des documents. Il faut donc verser la documentation et les déclarations en même temps que les documents auxquels elles se rapportent puisqu'elles constituent

les seules preuves attestant des mesures prises pour assurer l'intégrité des documents durant leur stade actif et semi-actif.

Une autre façon d'assurer le lien entre une partie de la documentation et les documents numérisés consiste à inscrire des renseignements sur le processus de numérisation dans les propriétés du fichier numérique, ou encore dans un autre fichier ou une fiche de métadonnées liée aux documents.

Lors de la réalisation d'une opération de numérisation, il est possible de rédiger une déclaration de numérisation en vue de documenter les mesures prises pour assurer l'intégrité des documents numériques. Cette déclaration, même si elle n'est pas obligatoire, peut s'avérer très utile dans le cas des documents destinés à être conservés sur une longue période. On trouve un modèle de déclaration de numérisation à l'annexe 2.

3.2 Organisation, identification et classification des fichiers

La création de fichiers numériques entraîne le recours à des outils informatiques appropriés pour la gestion de ces documents et à un ensemble de métadonnées. L'utilisation d'un système de gestion intégrée des documents est la voie à privilégier.

3.2.1 Classification et classement

La théorie archivistique précise que tous les documents doivent être classés dans des dossiers, eux-mêmes rattachés à des rubriques du plan de classification de l'organisme. Cela vaut autant pour les documents analogiques et numériques, que pour les documents numérisés au cours de leur cycle de vie.

Le même plan de classification doit servir à classer l'ensemble des documents d'un organisme, que celui-ci fasse partie d'un système de gestion intégrée des documents (GID) ou, en l'absence d'un tel système, qu'il soit reproduit sur un disque en réseau. De plus, pour une saine gestion documentaire, il est primordial d'établir un lien entre les documents analogiques et les documents numériques constituant un seul et même dossier « hybride ». L'utilisation d'un logiciel de gestion intégrée des documents facilite la gestion de ce type de dossiers, à l'aide de métadonnées spécifiques pour les dossiers et les documents, en plus de rendre possible l'application des mêmes règles d'accès aux documents.

3.2.2 Identification, repérage et métadonnées

Pour assurer le repérage des documents, il est nécessaire de les identifier d'une manière détaillée et significative. Cela est particulièrement vrai pour les documents numériques. En effet, il est plus fastidieux d'ouvrir chaque document numérique d'un dossier pour en connaître le contenu que de passer en revue les documents qui se trouvent dans une chemise papier. L'ajout d'éléments d'identification (métadonnées) est donc essentiel dès la création des documents. Cela facilitera l'identification, le repérage et l'utilisation des documents, ainsi que le maintien de leur intégrité. Ces ajouts peuvent se faire directement dans le fichier ou sur une fiche de métadonnées, qui est liée au dossier ou au document.

Pour obtenir de plus amples renseignements sur les métadonnées à ajouter aux dossiers et aux documents, veuillez consulter les profils de métadonnées gouvernementaux disponibles sur le portail de BAnQ, dans la section : *Accueil / Services / Services spécialisés en archivistique et gestion documentaire / CRGGID / Outils documentaires Outils documentaires*, à l'adresse suivante : http://www.banq.qc.ca/documents/services/archivistique_ged/Profils_metadonnees_gouvernementaux_Dos_Doc_version2_0_avril2009_VF.pdf

En outre, il importe d'identifier de manière normalisée et significative les fichiers informatiques des documents. Il est recommandé de nommer les fichiers en suivant une politique ou une directive institutionnelle, qui précise les éléments d'information à inclure et leur ordre de présentation. Par exemple, il faut déterminer le nombre de caractères et de mots permis, l'ordre des éléments, les symboles autorisés ou proscrits, etc. Il existe quelques modèles, mais l'important est de s'assurer que tous les employés d'une même organisation nomment leurs dossiers et leurs documents en respectant les mêmes règles.

3.3 Conservation et accès aux documents

Après avoir identifié et classé les documents numérisés, il faut les conserver de manière à en assurer la pérennité. Des informations détaillées sur les formats et supports de conservation sont présentées dans la section *4- Exigences techniques*.

Il est également essentiel de s'assurer que l'accès aux documents numériques est conforme aux règles et procédures en vigueur dans l'organisme pour la gestion documentaire. Cela veut dire qu'il doit exister des mécanismes de gestion et de contrôle des accès aux fichiers et répertoires, de façon à ce qu'une personne qui, par exemple, ne possède pas les droits d'accès au document papier ne puisse non plus avoir accès à la version numérisée.

3.4 Destruction sécuritaire des documents papier

La numérisation de documents source devant ensuite être détruits nécessite de prendre des dispositions pour s'assurer que la destruction se fasse de façon sécuritaire, et ce, afin de protéger les renseignements personnels et confidentiels, s'il y a lieu.

3.4.1 Méthodes de destruction sécuritaire

Il faut adopter des méthodes de destruction qui permettent de s'assurer qu'il est impossible, ou du moins extrêmement difficile, de reconstituer le document détruit.

Le déchiquetage est un procédé simple et efficace qui consiste à réduire le papier en fragments. C'est cette méthode qui est la plus couramment utilisée, et c'est également celle que recommande la Commission d'accès à l'information pour détruire l'information sur support papier. Il faut s'assurer que la taille des fragments de papier rende impossible la récupération de l'information.

La catégorisation de l'information permet d'identifier le degré de sensibilité des renseignements personnels ou confidentiels consignés sur le support papier. Pratique recommandée dans le réseau gouvernemental, la catégorisation est une forme de classification des actifs informationnels qui permet d'analyser les risques dans le cadre de la gestion de la sécurité. Les documents contenant des informations confidentielles sont identifiés avec une grande précision quant au degré de sensibilité de l'information à protéger. Le *Guide de destruction sécuritaire de l'information* (en cours d'élaboration par le ministère des Services gouvernementaux) fournit des indications précises sur les spécifications techniques et le pouvoir destructeur des déchiqueteurs qu'on peut utiliser pour répondre à ces exigences.

Le déchiquetage peut être effectué à l'interne ou être confié à des fournisseurs externes.

Le recyclage du papier non déchiqueté est une méthode de disposition des documents qui ne convient pas aux documents contenant des informations personnelles ou confidentielles. Le personnel doit être sensibilisé aux risques que présente cette méthode pour la sécurité de l'information.

3.4.2 Pratiques sécuritaires

Les pratiques élémentaires de sécurité physique ne doivent pas être négligées. Ainsi, les locaux où sont entreposés les documents papier en attente de destruction doivent être verrouillés et leur accès restreint aux personnes autorisées.

Il est recommandé de mettre en œuvre une politique de destruction des documents contenant des renseignements personnels et d'en surveiller l'application, d'élaborer des procédures de destruction et de définir les responsabilités. On doit tenir un registre indiquant la date et le procédé de destruction utilisé, le nom de la personne responsable de la destruction, ou encore, si la tâche est confiée à un tiers, le lieu de la destruction et le nom de la personne responsable du transport des documents.

Il est important de s'assurer que les fournisseurs externes et leurs sous-traitants, le cas échéant, s'engagent par contrat à respecter la confidentialité et à assurer la sécurité de l'information qui leur est confiée. La norme *ISO 15489*, qui s'applique à la gestion documentaire, de même que la Commission d'accès à l'information recommandent que soit exigée du fournisseur une attestation de destruction sécuritaire de l'information, indiquant la date, le lieu et le procédé de destruction utilisé, ainsi que les noms et signatures des personnes responsables ou témoins de la destruction. Enfin, le transport et la destruction des documents papier doivent, dans la mesure du possible, être réalisés en présence d'un représentant de l'organisme.

4. Exigences techniques

Dans le cadre d'un projet de numérisation, les exigences techniques recouvrent les éléments suivants : les outils de numérisation; la résolution des images, les formats de conservation des fichiers; les supports de conservation et de protection des fichiers.

4.1 Les outils de numérisation

Pour réaliser un projet de numérisation, certains outils sont essentiels : ordinateur, carte graphique, écran, numériseur, unité de mémoire, logiciel de numérisation. Pour que ces outils soient fonctionnels et qu'ils permettent le bon déroulement de l'opération, il importe de :

- s'assurer que les outils utilisés sont compatibles avec les diverses constituantes du processus;
- choisir du matériel standard de qualité, largement répandu et offrant un bon rapport qualité-prix. Le matériel doit correspondre aux normes et aux objectifs prévus et le fournisseur doit offrir un soutien technique;
- recourir à des outils de calibration tout au long de la chaîne de numérisation;
- prévoir l'adoption et l'application d'un plan de mises à jour et de conversions des fichiers et apporter des améliorations à intervalles réguliers afin que l'information numérisée soit toujours accessible;
- s'assurer que le personnel est bien formé et mettre sur pied un programme de formation, au besoin;
- s'assurer du degré de sécurité du système et des données.

4.2 Description des outils de numérisation

4.2.1 Ordinateur

Les ordinateurs et leurs interconnexions sont au cœur de tout projet de numérisation. Pour les projets d'envergure, il faut disposer d'ordinateurs réservés uniquement au travail de numérisation.

Les caractéristiques de ces ordinateurs (mémoire vive, processeur, système d'exploitation, capacité de stockage et d'interfaçage, cartes incluses, qualité des disques durs, etc.) doivent répondre aux exigences des applications qui seront utilisées à court et moyen terme.

Il faut aussi inventorier les périphériques et logiciels utiles au projet et connaître leurs caractéristiques. Les fabricants de logiciels et de périphériques indiquent les caractéristiques minimales que doit avoir un ordinateur pour faire fonctionner leurs produits. Il est généralement préférable de choisir un ordinateur offrant des capacités supérieures aux besoins actuels, surtout en ce qui concerne la mémoire vive et la vitesse du processeur.

S'il faut utiliser les ordinateurs déjà en place, il est important de porter une attention particulière aux spécifications des logiciels et périphériques qu'on veut acquérir. Il faut considérer notamment les points suivants :

- le système d'exploitation;
- la quantité de mémoire vive et la rapidité du processeur;
- l'espace disque disponible;
- les caractéristiques techniques du moniteur permettant un affichage d'images de qualité professionnelle;
- les prises et les cartes déjà installées permettant de communiquer avec les connexions des nouveaux périphériques (ex. : USB 2.0).

4.2.2 Écran à haute résolution et carte vidéo

L'écran est une composante essentielle de validation du processus de numérisation de textes et d'images. Il devrait mesurer au moins 19 pouces pour réduire la fatigue oculaire du personnel affecté à la numérisation. Il faut également porter une attention particulière au réglage de la résolution, au calibrage et à la propreté de l'écran.

Caractéristiques déterminant la qualité des images affichées sur écran :

- un affichage d'au moins 24 bits (16,8 millions de couleurs)
- une carte vidéo dotée d'une mémoire suffisante et optimisée pour le modèle de l'écran;
- une résolution égale ou supérieure à la résolution de numérisation. Pour un moniteur de 19 pouces, les résolutions optimales sont de 1024 x 768 et de 1280 x 1024;
- des possibilités de réglage de la luminosité, du contraste et des couleurs.

La qualité de l'affichage d'un écran dépend également de la performance de la carte graphique. Pour les projets d'envergure concernant des documents en couleur, il peut s'avérer utile de se procurer une carte graphique de meilleure qualité que celle de série qui est vendue avec l'ordinateur.

4.2.3 Numériseur

Il existe de nombreux types et modèles de numériseurs. Il faut évaluer les caractéristiques en fonction des besoins : volume de pages à numériser, alimentation automatique, recto verso, etc.

Sont aussi à prendre en considération :

- la résolution optique en points par pouce (ppp ou dpi);
- le spectre du numériseur pouvant rendre toutes les couleurs du document original;
- le traitement des petits formats, formats hors norme, formats continus, grands formats, copies carbonées, papiers épais, etc.;
- le comptage automatique des pages;
- le cycle de vie de l'appareil et son soutien technique.

Un seul numériseur ne suffit pas toujours à répondre à l'ensemble des besoins. Selon les types de documents, les formats et la fréquence d'utilisation de l'appareil, il peut être nécessaire, lorsque la numérisation se fait à l'interne, d'acquérir plus d'un numériseur. Il est également possible de faire numériser une partie des documents à l'externe.

Principaux types de numériseurs : *Numériseur à plat*

Le numériseur à plat est le plus répandu des appareils de saisie d'images. Il convient à de nombreuses utilisations, il est facile à utiliser et largement disponible. Il permet de saisir du texte et des images non transparentes, ainsi que des objets transparents comme les diapositives 35 mm. Si l'on a un grand nombre de diapositives à numériser, il est toutefois préférable d'acquérir un appareil spécialisé ou de recourir aux services d'une firme externe.

Tous les numériseurs de cette catégorie utilisent la même technologie de base, selon laquelle un capteur lumineux et une source lumineuse montés sur un bras mobile balayent le document immobile placé sur une plaque en verre. Des alimenteurs automatiques sont disponibles sur certains modèles et peuvent augmenter le rythme de production tout en réduisant la fatigue de l'opérateur lors de la manipulation d'ensemble de documents uniformes.

Numériseur avec caméra planétaire ou appareil photo numérique

Pour les documents reliés et les grands formats, il est possible d'utiliser un numériseur avec caméra planétaire ou un appareil photo numérique. Dans ces deux cas, l'appareil et son support sont réglés pour ne pas endommager les documents fragiles. Plusieurs accessoires sont en outre disponibles pour réduire les risques d'endommagement.

Les appareils photo numériques permettent de produire à faible coût des images de bonne qualité pour les formats courants, sans nécessiter des installations fixes comme les caméras planétaires. Il faut toutefois s'assurer que le personnel ait de bonnes connaissances de base en photographie numérique.

Autres types de numériseur

Il existe d'autres types de numériseurs pour les microfilms, les documents reliés et les grands formats. À moins que le volume des documents à numériser soit très important, il est généralement préférable de confier la numérisation à un fournisseur externe, l'emploi de ces appareils exigeant un savoir-faire qui s'acquiert par une longue expérience.

4.2.4 Logiciel de numérisation

Il importe de choisir le logiciel de numérisation en fonction des besoins du projet de numérisation, du budget, des normes à respecter et des possibilités et options offertes. La plupart des numériseurs sont fournis avec un logiciel de numérisation généralement adéquat pour les besoins courants, y compris la reconnaissance optique des caractères et le traitement sommaire d'images. On peut aussi acquérir séparément des logiciels spécialisés pour répondre à des besoins spécifiques.

Caractéristiques à considérer lors du choix du logiciel :

- sa fiabilité ou sa maturité technologique;
- sa convivialité;
- les possibilités de l'adapter à des besoins spécifiques;
- les formats offerts pour l'importation et l'exportation : PDF, PDF/A-1, TIFF, JPEG, JPEG 2000, etc.
- les fonctions de traitement par lots et de macros pour des opérations répétitives;
- les options et la fiabilité du module de reconnaissance optique de caractères (OCR);
- l'enregistrement de l'espace de travail afin de reprendre un travail en cours;

- les fonctions intégrées de gestion d'images ou de métadonnées;
- le réglage de la luminosité, du contraste, des blancs, de l'échelle de gris, des couleurs et de la plage dynamique.

4.2.5 Logiciel de reconnaissance optique des caractères

Les numériseurs sont souvent fournis avec un logiciel permettant d'appliquer une reconnaissance optique des caractères (OCR) aux documents textuels (sauf les manuscrits). Si le volume des documents à traiter est important, il est bon de s'assurer que le logiciel offre la possibilité de suggérer des corrections pour les mots qu'il n'a pu reconnaître, en plus de pouvoir traiter les textes disposés en plusieurs colonnes.

4.3 Résolution et mode de l'image

La résolution détermine largement la qualité de l'image numérisée. Elle se mesure en points par pouce (ppp ou dpi). Plus la résolution est élevée, meilleure est la définition de l'image, mais cette qualité se paie au prix de fichiers de plus grande taille. Il est donc important d'évaluer les besoins pour déterminer la résolution appropriée. Très souvent, une résolution très élevée n'apporte rien à la qualité perceptible des images et constitue un gaspillage de temps (pour la numérisation) et d'espace-disque (pour le stockage des fichiers).

Le tableau ci-dessous résume les résolutions d'usage courant pour les différents types de documents.

Support d'origine	Dimensions	Résolution
Documents textuels sur papier	Tous formats	300 ppp
Documents textuels sur microfilms		1200 ppp et plus
Documents graphiques sur papier (photos, cartes, plans)	Jusqu'à 8,5 x 11 Plus de 8,5 x 11 po	600 ppp 300 ppp

Dans le cas des documents textuels et graphiques sur papier, une résolution de 300 ppp est suffisante pour reproduire les documents à leur taille originale. Pour obtenir des agrandissements, il faut choisir une résolution supérieure (par exemple, 600 ppp, qui multiplie par quatre la taille des fichiers).

Un autre élément important à prendre en considération est le mode colorimétrique, le type d'organisation des couleurs de l'image. Plus le mode offre une gamme élevée de couleurs, plus les fichiers sont volumineux.

- Le mode noir et blanc (bitonal, 1 bit) restitue approximativement l'original et n'est pas recommandé, même s'il est le plus avantageux du point de vue de la taille des fichiers;
- Le mode tons de gris (8 bits) offre 256 tons de gris. Il est recommandé pour la plupart des documents monochromes, textuels ou photographiques. Noter qu'on peut aussi numériser en couleurs 8 bits, ce qui est généralement suffisant pour les documents textuels comportant quelques couleurs (en-têtes, signatures, etc.);
- Le mode vraies couleurs (24 bits) offre 16 millions de couleurs et s'impose pour les documents de type photographique en couleur.

4.4 Les formats de conservation des fichiers

Il existe un grand nombre de formats de fichiers. Très peu sont recommandés pour la conservation à long terme des fichiers.

BAnQ recommande présentement les trois formats suivants pour la conservation de documents numérisés : TIFF, PDF/A-1 et XML.

4.4.1 TIFF non compressé

Le format TIFF (Tagged Image File Format) est le format le plus utilisé jusqu'à présent pour la conservation d'images numériques. Bien que la taille des fichiers résultants soit considérable, le format TIFF est généralement retenu pour la conservation à long terme des images numériques.

4.4.2 PDF/A-1

Le format ouvert PDF/A-1 est défini par la norme ISO 19005-1⁹. Il s'agit d'un format image adapté aux documents textuels. Il préserve la mise en page, les polices de caractères et la mise en forme, ce qui le rend attrayant pour la conservation et la diffusion. De plus, il s'agit d'un standard ouvert et libre de droits.

⁹ Gestion de documents - Format de fichier des documents électroniques pour une conservation à long terme - Partie 1 : Utilisation du PDF 1.4 (PDF/A-1) ICS 35.240.30; 37.100.99.

Le format PDF/A comporte deux niveaux de conformité, le PDF/A 1a et le PDF/A 1b. Le niveau 1a concerne la conformité de la structure logique du document avec la norme, alors que le niveau 1b se limite à la conformité visuelle du document.

Pour les documents simplement numérisés en mode image (sans OCR), le niveau 1b s'applique. On retient le niveau 1a pour la conservation des documents structurés (p. ex, les fichiers produits par les logiciels de traitement de texte ou numérisés en mode image avec OCR). Dans le doute, BAnQ suggère de se conformer au niveau de conformité PDF/A 1a comme il englobe le niveau 1b.

4.4.3 XML

Le format normalisé XML produit des fichiers structurés basés sur un langage balisé favorisant l'échange de données. Il offre de très grandes possibilités pour l'utilisation des métadonnées nécessaires à la gestion des documents.

Le format XML est également intéressant pour la conservation à long terme des documents puisqu'il sépare la présentation et le contenu des documents. De plus, il s'agit d'un standard ouvert et libre de droits.

BAnQ recommande l'utilisation du XML ou de formats basés sur l'utilisation de schéma XML comme *l'Open Document Format (.odf)* et *OpenXML (.docs)*. Ces formats sont particulièrement appropriés pour la création et la conservation des documents de type traitement de texte nés numériques ou qui résultent d'une reconnaissance optique de caractères (OCR).

4.5 Les supports de stockage et la protection des fichiers

Il existe deux principaux types de supports de stockage qui permettent de conserver les fichiers, à savoir les supports optiques et les supports magnétiques.

4.5.1 Supports optiques

Les supports optiques les plus connus sont les CD-ROM (Compact Disc ou disques optiques compacts) et les DVD (Digital Versatile Disc).

Les CD-ROM ont une capacité de stockage six fois plus faible que les DVD et ne sont pas recommandés pour les projets de numérisation d'envergure. Les DVD, pour leur part, offrent des capacités de stockage plus considérable (4,7 Go pour un DVD standard simple couche). Malgré leur capacité de stockage accrue, il vaut mieux éviter les DVD double couche, beaucoup moins répandus que les DVD simple couche.

Il faut éviter les disques réinscriptibles DVD-RW (DVD Rewritable) et choisir des disques inscriptibles (DVD-R : DVD Recordable) de qualité « Archives » (par ex. *Verbatim Ultralife Gold* ou *Mitsui/MAM-A Gold*). Ces disques sont plus coûteux que les DVD-R couramment disponibles sur le marché mais ont une espérance de vie plus longue. On les obtient auprès de fournisseurs spécialisés.

Pour graver sur les DVD, il faut un graveur de DVD et un logiciel de gravure. Ils sont généralement fournis avec les ordinateurs actuels, mais on peut aussi acquérir des logiciels spécialisés offrant davantage de possibilités.

Pour assurer la longévité des supports optiques et l'accès à long terme aux informations qu'ils contiennent, certaines précautions s'imposent :

- éviter de graver la capacité maximale de données sur un disque, afin que les données ne se retrouvent pas à la périphérie du disque, qui est une zone plus fragile car elle est exposée aux traces de doigts;
- graver les fichiers à la vitesse recommandée par le fabricant des disques : une vitesse supérieure peut entraîner la perte de données;
- lors de la gravure, fermer tous les programmes et désactiver toutes les applications sur l'ordinateur qui sert à la gravure;
- conserver les disques à des conditions appropriées (l'idéal est une température de 20 °C et une humidité relative de 40 %);
- graver au moins deux exemplaires de tous les disques et les conserver dans des lieux différents, si possible des édifices différents;
- ranger les disques verticalement pour éviter le gauchissement;
- utiliser des boîtiers conçus pour la conservation des supports optiques;

- éviter d'exposer les disques au soleil, à la chaleur ou à l'eau;
- éviter les autocollants sur les disques; inscrire des renseignements uniquement sur la partie centrale et transparente du disque, qui ne comporte pas de données;
- effectuer une vérification des disques une fois la gravure terminée : certains logiciels de gravure peuvent faire cette vérification automatiquement;
- prévoir des vérifications périodiques (ex. : tous les trois à cinq ans).

4.5.2 Supports magnétiques

Les supports magnétiques les plus répandus sont les disques rigides et les cartouches magnétiques. Leurs capacités de stockage sont considérables et s'accroissent sans cesse.

Les disques magnétiques peuvent résider sur des serveurs situés à l'interne ou à l'externe. Le temps d'accès aux données et le taux de transfert sont plus rapides que pour les supports optiques. Un protocole de sécurisation des données de type *RAID* (Redundant Array of Independent or Inexpensive Discs) doit être employé pour assurer la protection des données en cas de panne. Il faut en outre assurer très régulièrement la prise de copies de sauvegarde des fichiers et conserver ces copies à l'extérieur de l'organisation.

Les cartouches magnétiques se présentent sous des formes très diverses (*DLT, DAT, LTO, etc.*). (*Digital Linear Tape : DLT*). Elles sont inutilisables pour l'accès courant aux documents qui y sont stockés mais offrent une grande capacité de stockage à peu de frais. Ici encore, il faut en conserver plus d'une copie, dans des lieux différents, et vérifier périodiquement leur état, en recopiant au besoin les données sur un autre support.

4.5.3 Choix de formats et de supports de conservation

Dans le monde des technologies de l'information, en continuelle évolution, aucun format ou support ne peut assurer une pérennité absolue. Il est préférable de s'en tenir aux formats et aux supports les plus répandus, qui offrent la meilleure garantie de longévité.

Il faut aussi considérer qu'à long terme, des transferts et migrations de formats et de supports s'imposeront nécessairement pour assurer l'intégrité et la permanence de l'accès aux informations.

5. Aide-mémoire pour la numérisation de documents

Dans le cadre d'un projet de numérisation, il est important de respecter les obligations légales, désignées par un astérisque (*), de même que les recommandations de BAnQ, que ce soit en matière d'élaboration ou de modification du calendrier de conservation ou, par exemple, du choix de format des documents transférés sur un nouveau support.

Cette section, présentée sous forme de tableau, peut être utilisée comme aide-mémoire.

5.1 Cadre juridique - obligations légales et recommandations

Documenter le processus de numérisation

- Documentation à conserver tout au long du cycle de vie du document. On y note le format du document source, le procédé de transfert, les garanties offertes, etc. *

Modifier et faire approuver le calendrier de conservation*

- Approbation à obtenir avant la destruction des documents source, préférablement avant d'entreprendre la numérisation **D**

Pour détruire un document source, il faut :*

- documenter le transfert et valider la conformité du résultat du transfert
- assurer la protection des renseignements personnels et confidentiels lors de la destruction
- agir conformément au calendrier de conservation

Libérer les droits d'auteur, s'il y a lieu*

Classifier les documents selon le plan de classification de l'organisme afin d'en assurer l'identification et le repérage*

Limiter l'accès aux renseignements personnels et confidentiels par des règles d'habilitation et des mécanismes de contrôle d'accès aux documents et aux installations de numérisation*

Utiliser la *Déclaration de numérisation d'archives*

5.2 Projet de numérisation - recommandations

- Dresser les objectifs et les buts du projet de numérisation ○
- Déterminer les documents à numériser et l'échéancier du projet □
- Effectuer une analyse des besoins pour le projet et dresser l'inventaire des ressources disponibles à l'interne (matérielles, humaines et financières) □
- Respecter les étapes de la chaîne de numérisation (voir pages 11-14) ┌-┐
u

5.3 Exigences archivistiques - recommandations

- Utiliser un système de gestion intégrée des documents, afin de gérer les documents sur papier et les documents numériques de la même manière ○
- Classer les documents papier et numériques selon le même plan de classification ┌-┐
- Identifier les documents de façon détaillée et significative □
- Ajouter des métadonnées aux documents numériques dès leur création ┌-┐
└┘
- Adopter des méthodes de destruction adéquates qui assurent la protection des renseignements personnels et confidentiels □
 - Déchiquetage à privilégier

5.4 Exigences techniques - recommandations

- S'assurer d'avoir et d'utiliser le matériel et les outils adéquats Q
- Choisir le support en fonction de la quantité de documents, la rapidité d'accès aux documents, les coûts et les besoins de sécurité ┌-┐
└┘

Choisir des formats et des résolutions favorisant la conservation permanente des documents numériques :

Support d'origine	Dimensions	Résolution	Format
Documents textuels sur papier	Tous formats	300 ppp	TIFF PDF/A XML
Documents textuels sur microfilms		1200 dpi et plus	TIFF
Documents graphiques sur papier (photos, cartes, plans)	Jusqu'à 8,5 x 11 po Plus de 8,5 x 11 po	600 dpi 300 dpi	TIFF PDF/A

Bibliographie

De la théorie à la pratique - Didacticiel d'imagerie numérique, Cornell University,
<http://www.library.cornell.edu/preservation/tutorial-french/contents.html>
(site consulté le 2 octobre 2009).

Dématérialisation et archivage électronique - Mise en œuvre de l'ILM (Information Lifecycle Management), par Jean-Marc Rietsch et al., Paris, Dunod, 2006, 207 p.

Document et numérisation : enjeux techniques, économiques, culturels et sociaux,
par Jacques Chaumier, Paris, ADBS, 2006, 119 p.

Gestion de documents - Format de fichier des documents électroniques pour une conservation à long terme - Partie 1 : Utilisation du PDF 1.4 (PDF/A-1) ICS
35.240.30; 37.100.99.

Guide d'imagerie numérique - Numérisation de l'information consignée sur des supports traditionnels, par M. Bélanger, J. M. Demers et C. Minotto, Archives nationales du Québec, 1998 (version corrigée).

Guide relatif à la catégorisation des documents technologiques en matière de sécurité, ministère des Services gouvernementaux, 2003, (SGQRI 036).

Guide de destruction sécuritaire de l'information, ministère des Services gouvernementaux, janvier 2009 (version 1.0).

Guide pour la destruction des documents renfermant des informations personnelles, Commission d'accès à l'information, s.d.
http://www.cai.gouv.qc.ca/06_documentation/01_pdf/destruct.pdf

ISO 15489 -2001 - Information et documentation - Records management.

La numérisation : un guide pratique, Bibliothèque et Archives nationales du Québec, 2007 (présentation Power Point dans le cadre du pré-congrès de l'Association des archivistes du Québec, le 31 mai 2007).

Numérisation de documents inscrits au calendrier de conservation : questionnaire, Archives nationales du Québec, 2003.

Orientations de BANQ en matière de numérisation pour les organismes publics, par C. Otis et J. Nadeau, Bibliothèque et Archives nationales du Québec, 2007 (présentation Power Point).

PDF in a Nutshell: Long Term Archiving with PDF, par Olaf Drümmer, Alexandra Oettler und Dietrich von Seggern, Berlin, PDF/A Competence Center, 2007, 87 p.

Recommandations relatives à la gravure, à la conservation et à l'évaluation des CD-R, Paris, Direction des archives de France, mars 2005, 20 p.

Annexe 1 - Déclaration de numérisation d'archives

DÉCLARATION DE NUMÉRISATION D'ARCHIVES

ATTENTION : *La Loi concernant le cadre juridique des technologies de l'information (L.R.Q. c. C-1.1) n'exige pas la signature d'une déclaration de numérisation d'archives pour qu'en soit préservée sa valeur juridique. BAnQ propose cependant Ce présent modèle aux organismes qui choisissent d'utiliser un tel outil*

Nom et adresse de
l'organisme détenteur des
archives

Nom et adresse de l'unité
administrative ou de l'entreprise
responsable de la numérisation

JE
SOUSSIGNÉ(E)

exerçant mes fonctions dans

l'unité administrative ou dans l'entreprise responsable de la numérisation de ces archives :

1° **DÉCLARE** que j'ai été formellement habilité(e) par l'organisme détenteur mentionné ci-dessus à numériser les archives identifiées au paragraphe 2°, en vertu du document suivant :

(Titre et date du document habilitant)

2° **GARANTIS** que j'ai numérisé systématiquement et sans lacune les archives suivantes :

Titre de la série
(tel qu'inscrit au calendrier, si applicable, ou
au
plan de classification du détenteur)

Renseignements descriptifs
(support(s), quantité, dates
extrêmes, numéros
séquentiels,

	caractères alphabétiques ou alphanumériques, etc. ¹⁰)

Cette série est décrite dans la règle de conservation approuvée n° _____ du calendrier de conservation de l'organisme détenteur des archives identifiées ci-dessus.

3° **DÉCLARE** que la numérisation du ou des documents mentionné(s) au paragraphe 2° a été effectuée dans les conditions suivantes :

Circonstances de la numérisation (dont le lieu)	
Date de la numérisation	
Moyen(s) utilisés pour assurer l'intégrité de la copie (dont le format utilisé, la résolution de la numérisation, la documentation produite)	

4° **CERTIFIE** que l'information contenue dans les documents numérisés est la même que celle contenue dans les documents sources, après avoir effectué les contrôles suivants de la qualité et de la quantité :

EN FOI DE QUOI, j'ai signé à _____, ce

jour du mois de _____ 20__ .

Signature

(Élément facultatif laissé à la discrétion de l'organisme détenteur des archives)

AFFIRMÉ SOLENNELLEMENT devant moi à _____, ce _____ jour de _____ 20__ ,
(Signature de la personne qui assermente et autres renseignements requis par la loi)

¹⁰. C'est-à-dire tout renseignement permettant d'identifier un groupe d'archives donné à l'intérieur d'une série plus large. Le ou les supports et la quantité précise d'archives numérisées devraient toujours être mentionnés.