

La numérisation des documents administratifs

Méthodes et recommandations

Version revue et corrigée

Juin 2010

*Bibliothèque
et Archives
nationales*

Québec 

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2010
ISBN 978-2-550-60182-1

Table des matières

Table des matières	3
Introduction	4
1. Cadre juridique.....	6
1.1 Inscription au calendrier de conservation	6
1.2 Intégrité et valeur juridique	7
1.3 Documentation du processus de numérisation.....	8
1.3.1 Déclaration de numérisation	9
1.3.2 Procédure de numérisation	9
1.3.3 Fichier de numérisation et métadonnées.....	10
1.4 Destruction et conservation des documents source	10
1.5 Respect du droit d'auteur.....	11
1.6 Accès à l'information	11
1.7 Protection des renseignements personnels et confidentiels.....	12
2. Projet de numérisation	13
2.1 Définition de projet : les objectifs recherchés	13
2.2 Choix des documents à numériser et échéancier du projet.....	13
2.3 Analyse des besoins et des ressources disponibles	14
2.4 Numérisation des documents	15
2.4.1 Gestion matérielle des documents.....	15
2.4.2 Réalisation de projets pilotes.....	16
2.4.3 Numérisation des documents	16
2.4.4 Contrôle de la quantité et de la qualité	16
3. Exigences archivistiques.....	19
3.1 Calendrier de conservation.....	19
3.1.1 Mise à jour.....	19
3.1.2 Pratiques sécuritaires	22
4. Exigences techniques	23
4.1 Les outils de numérisation	23
4.1.1 Ordinateur.....	23
4.1.2 Écran à haute résolution et carte vidéo	24
4.1.3 Numériseur	24
4.1.4 Logiciel de numérisation et de reconnaissance optique de caractères.....	26
4.2 Résolution et mode de l'image	26
4.3 Formats de conservation des fichiers et compression.....	27
4.3.1 TIFF non compressé	27
4.3.2 PDF/A-1.....	28
4.3.3 XML.....	28
4.3.4 La compression	28
4.4 Supports de stockage et protection des fichiers	29
4.4.1 Supports optiques.....	29
4.4.2 Supports magnétiques.....	30
4.4.3 Choix de formats et de supports de conservation	30
Annexe : Aide-mémoire pour la numérisation de documents.....	31
Bibliographie	33

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 de l'équipement et la réduction de ses 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 publics du Québec. Il vise à leur fournir 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*.

Ce 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 à certaines conditions, l'élimination des documents source sur supports papier ou microfilm.

Ce guide compte quatre sections, celles-ci couvrant les divers volets d'une opération de numérisation dans un cadre de gestion documentaire, ainsi qu'un aide-mémoire. 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 efficace 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.

Ce guide comporte aussi en annexe un aide-mémoire qui présente une synthèse des obligations juridiques ainsi que les principales recommandations de BAnQ applicables aux projets de numérisation et de destruction des documents source analogiques. Il peut être utilisé par les organismes comme une liste de validation lors de la réalisation d'un projet de

numérisation pour s'assurer de ne pas oublier d'étapes ou d'éléments à prendre en considération.

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 être utile à toutes les 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 juridiques qui s’y appliquent et les respecter. Dans cette partie, sont présentées les principales composantes du cadre juridique dont il faut tenir compte.

1.1 Inscription au calendrier de conservation¹

La 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;
- de 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 source 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 source comprenne la période requise pour effectuer un contrôle de qualité exhaustif, cette période devant être indiquée 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								

1. Loi sur les archives (L.R.Q., c. A-21.1, art. 7 et 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 œuvre 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 adéquate 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 l'information est libre. Le support est ainsi une caractéristique neutre par rapport à la valeur juridique du document.

2. *Loi concernant le cadre juridique des technologies de l'information* (L.R.Q., c. C-1.1, art. 5, 6, 9, 10, 12, 14 et 68).

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 *LCCJTI* 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.

La documentation peut prendre plusieurs formes selon le contexte, la fréquence et les objectifs poursuivis par la numérisation. Elle peut, par exemple, prendre celle d'une déclaration de numérisation regroupant l'ensemble des renseignements de documentation ou celle d'une procédure de numérisation complétée par une série de métadonnées ou des fichiers de compilation. La première s'avère plus appropriée à une opération de numérisation de séries homogènes et définie dans le temps, alors que la seconde est plus appropriée pour de la numérisation à la pièce et par de

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).

multiples utilisateurs ou pour un organisme qui se dote d'un service permanent de numérisation. Ces deux méthodes peuvent être simultanément en usage dans un organisme, selon les besoins. Par exemple, il sera plus simple pour un utilisateur qui numérise un document par jour de lui ajouter les données de contexte comme métadonnées alors que, pour un utilisateur qui numérise plusieurs documents en série, il pourrait être plus simple de recourir à un fichier de numérisation.

1.3.1 Déclaration de numérisation

Une déclaration peut être utilisée lors de la numérisation pour une quantité donnée pendant une période donnée ou lors d'une activité récurrente (exemple : numérisation d'une série de dossiers de clients pendant la période estivale).

Dans le cas où un organisme choisit d'utiliser une déclaration de numérisation établie selon le modèle d'une déclaration appuyant une copie à l'intention d'un tribunal, celle-ci peut comprendre les éléments suivants :

- le nom et l'adresse de l'organisme détenteur des archives;
- le nom et l'adresse de l'unité administrative ou de l'entreprise responsable de la numérisation;
- la signature de la personne qui exerce ses fonctions dans l'unité administrative ou dans l'entreprise responsable de la numérisation de ces archives;
- le titre du document qui habilite la personne à numériser les documents (exemples : résolution du conseil d'administration ou contrat de numérisation pour une firme externe);
- le titre de la série (tel qu'inscrit au calendrier de conservation ou au plan de classification), le support, la quantité et les dates extrêmes;
- le numéro de règle de conservation;
- le lieu de la numérisation;
- la date de la numérisation;
- le format de numérisation, la résolution utilisée et la documentation produite;
- le contrôle de la qualité et de la quantité effectué;
- la signature;
- l'assermentation.

1.3.2 Procédure de numérisation

Une procédure de numérisation peut être utilisée quand la numérisation est effectuée à la pièce par le personnel de l'organisme ou que celui-ci possède une équipe et une infrastructure consacrées à la numérisation. La procédure doit faire l'objet d'une approbation par les instances décisionnelles de l'organisme et il est primordial que le personnel se conforme aux exigences qui y sont énoncées. Elle a l'avantage d'alléger le processus et les tâches des individus.

Une procédure de numérisation doit minimalement contenir les éléments suivants :

- le nom des personnes ou des groupes de personnes habilitées à numériser;
- le procédé de transfert (numérisation) visé par la procédure;
- les mesures à prendre en vue de la préparation physique des dossiers et des documents avant le début d'une opération de numérisation;
- les politiques et mécanismes à respecter en ce qui concerne le contrôle de l'accès;
- les éléments du contrôle de la qualité et de la quantité à effectuer afin de pouvoir attester l'intégrité des documents numérisés;
- les formats de numérisation, les résolutions à utiliser et la documentation afférente produite (registre, fichier, etc.);
- le matériel utilisé et les garanties offertes par celui-ci en matière d'intégrité.

Une procédure de numérisation ne peut à elle seule constituer toute la documentation du processus de numérisation et assurer l'intégrité d'un document numérisé. Selon le contexte et les besoins, elle peut être accompagnée d'une déclaration simplifiée, d'un fichier de numérisation ou d'un registre de numérisation dans lequel on trouve le nom de la personne qui a réalisé la numérisation, la date de ce travail, la résolution ainsi que la liste des documents numérisés. Elle peut également être complétée par l'ajout direct de métadonnées au document numérisé.

1.3.3 Fichier de numérisation et métadonnées

En plus de connaître la nature du matériel utilisé et du processus de numérisation, il est important d'être en mesure d'identifier ce qui a été numérisé, quand, comment et par qui. La procédure de numérisation traite donc du général alors que les fichiers de numérisation portent sur le particulier.

Lors de l'utilisation d'un fichier de numérisation ou de l'ajout de métadonnées aux documents, il faut trouver minimalement les éléments suivants :

- le nom du fichier;
- le nom de la personne qui a effectué la numérisation;
- la date de la numérisation;
- la résolution de la numérisation.

1.4 Destruction et conservation des documents source⁵

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 conditions sont respectées.

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

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 peuvent devoir être conservés. Il s'agit des documents qui, sur leur support d'origine, possèdent une valeur archivistique, historique ou patrimoniale selon 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 un droit exclusif de l'auteur, il faut s'assurer, avant de procéder à la numérisation, de détenir les droits d'auteur sur les 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.

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, sur demande, 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.

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

7. *Loi 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 et 16.1).

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 et ne sont pas les renseignements personnels à protéger : par exemple, le nom d'une personne physique n'est pas considéré comme un renseignement personnel s'il n'est accompagné d'aucune 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 contrôler l'accès aux renseignements personnels et confidentiels, il est nécessaire, avant la numérisation, d'adopter des règles d'habilitation. De 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*.

8. *Loi 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 et 63.1).

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 numérisés 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 les 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 de l'équipement 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 confiée à un fournisseur externe. Il n'est pas rare de recourir simultanément à ces deux formules pour répondre à l'ensemble des besoins. Les principaux avantages de la numérisation à l'interne sont les suivants : meilleure accessibilité aux documents pendant le projet, réduction des risques liés à la manipulation des documents fragiles, meilleur contrôle de la qualité et 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 :

- Gestion 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 par la suite.

2.4.1 Gestion 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 de l'équipement. 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 l'équipement, 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.

Bien qu'il faille éviter l'utilisation des documents originaux pouvant être altérés par de multiples expositions à la lumière et des manipulations répétées lors de projets pilotes, l'utilisation des documents source est parfois nécessaire pour effectuer les essais. Dans de tels cas, il est impératif de prendre toutes les précautions pour assurer la préservation des documents fragiles.

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 degré 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 de respecter les règles d'accès et de sécurité et de manipuler 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 approprié 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 ligne n'est coupée, 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 valider plus facilement les résultats et de respecter les normes convenues. De plus, cela permet de déceler des erreurs récurrentes telles que :

- la perte de détails dans les points d'ombre 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 photo 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 des formes 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 signes 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 au préalable 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 / Espace professionnel / Archiviste / Services aux organismes publics* » sur le portail de BAnQ, à l'adresse suivante : www.banq.qc.ca/services/archivistique_ged/public/index.html.

3.1.2 Création de la documentation

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 :

- par la création d'une règle spécifique pour cette série documentaire;
- par leur ajout comme types de documents aux règles des documents originaux.

Dans les deux cas, il importe que le lien entre les documents numérisés, la documentation du processus de numérisation et, le cas échéant, 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, sur une fiche de métadonnées liée aux documents ou dans un fichier de numérisation.

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 nés 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 / Espace professionnel / Archiviste / Profils de métadonnées gouvernementaux* », à l'adresse suivante :
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 de gestion documentaire en vigueur dans l'organisme. 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 confidentiels (dont les renseignements personnels), 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 déchetage 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 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 renseignements confidentiels 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 confié à des fournisseurs externes.

Le recyclage du papier non déchiqueté est une méthode de destruction des documents qui ne convient pas aux documents contenant des renseignements confidentiels. 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 confidentiels 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 d'exiger 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 concernent les éléments suivants : les outils de numérisation; la résolution des images; les formats de conservation des fichiers; les supports de stockage 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, logiciel de numérisation. Pour que ces outils soient fonctionnels et permettent le bon déroulement de l'opération, il importe :

- de s'assurer que les outils utilisés sont compatibles avec les diverses constituantes du processus;
- de 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;
- de recourir à des outils de calibration tout au long de la chaîne de numérisation;
- de prévoir l'adoption et l'application d'un plan de mise à jour et de conversion des fichiers et d'apporter des améliorations à intervalles réguliers afin que l'information numérisée soit toujours accessible;
- de s'assurer que le personnel soit bien formé et de mettre sur pied un programme de formation, au besoin;
- de s'assurer du degré de sécurité du système et des données.

4.1.1 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 à 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).

4.1.2 Écran à haute résolution et carte vidéo

L'écran est une composante essentielle pour la vérification du résultat de la numérisation de textes et d'images. Il devrait mesurer au moins 19 pouces de large 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 à respecter pour assurer la qualité des images affichées :

- 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.

4.1.3 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, soit un capteur lumineux et une source lumineuse montés sur un bras mobile qui 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'ensembles 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 numériseurs

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.

4.1.4 Logiciel de numérisation et de reconnaissance optique de caractères

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.

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.

4.2 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 les formats	300 ppp
Documents textuels sur microfilm		1200 ppp et plus
Documents graphiques sur papier (photos, cartes, plans)	Jusqu'à 8,5 x 11 po 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. Ainsi :

- 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. Notez 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 photographies en couleurs.

4.3 Formats de conservation des fichiers et compression

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 à long terme de documents numérisés : TIFF, PDF/A-1 et XML.

4.3.1 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.

4.3.2 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.

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, le niveau 1b s'applique. Pour les documents numérisés en mode image dont on réalise ensuite la lecture optique (OCR), il faut se conformer au format PDF/A 1a.

On retient le niveau 1a pour la conservation des documents structurés (par exemple, les fichiers produits par les logiciels de bureautique comme Word, Excel, PowerPoint, OpenOffice, etc.). Si on utilise les versions les plus récentes de ces logiciels, on peut aussi envisager de les sauvegarder en format XML (voir ci-dessous).

4.3.3 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 (.docx). Ces formats sont particulièrement appropriés pour la création et la conservation des documents de type traitement de texte numériques ou qui résultent d'une reconnaissance optique de caractères (OCR).

4.3.4 La compression

Pour réduire la taille du résultat de la numérisation, il existe des formats de fichiers compressés. La plupart (ex. : JPEG ou TIFF avec compression G4) entraînent une dégradation des fichiers originaux. D'autres, plus rares (ex. : JPEG 2000), n'entraînent aucune perte. Les formats compressés avec perte ne doivent pas être utilisés comme format d'archivage, surtout quand il y a élimination des documents source. On peut toutefois y recourir pour produire

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.

une version de diffusion des fichiers.

Dans le cadre de projets de numérisation de documents administratifs, il est recommandé de créer des fichiers non compressés. Si le recours à un format de fichier compressé est nécessaire, il faut en choisir un n'entraînant aucune perte et n'étant soumis à aucun brevet.

4.4 Supports de stockage et 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.

4.4.1 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 considérables (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 » (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 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, une zone plus fragile car elle est exposée aux traces de doigts;
- graver les fichiers à la vitesse recommandée par le fabricant des disques puisqu'une vitesse supérieure de gravure peut entraîner la perte de données;
- lors de la gravure, fermer tous les programmes et désactiver toutes les applications sur l'ordinateur;

- 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 et 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).

4.4.2 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.

4.4.3 Choix de formats et de supports de conservation

Aucun format ou support ne peut assurer une pérennité absolue. Il est cependant 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.

Annexe : Aide-mémoire pour la numérisation de documents

Dans le cadre d'un projet de numérisation, il est important de respecter les obligations juridiques, 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.

Cadre juridique – Obligations juridiques et recommandations

Documenter le processus de numérisation <ul style="list-style-type: none"> 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.* 	<input type="checkbox"/>
Modifier et faire approuver le calendrier de conservation* <ul style="list-style-type: none"> Approbation à obtenir avant la destruction des documents source, préférablement avant d'entreprendre la numérisation 	<input type="checkbox"/>
Pour détruire un document source, il faut :* <ul style="list-style-type: none"> 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. 	<input type="checkbox"/>
Libérer les droits d'auteur, s'il y a lieu*	<input type="checkbox"/>
Classifier les documents selon le plan de classification de l'organisme afin d'en assurer l'identification et le repérage*	<input type="checkbox"/>
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*	<input type="checkbox"/>
Utiliser la <i>Déclaration de numérisation d'archives</i>	<input type="checkbox"/>

Projet de numérisation – Recommandations

Dresser les objectifs et les buts du projet de numérisation	<input type="checkbox"/>
Déterminer les documents à numériser et l'échéancier du projet	<input type="checkbox"/>
Effectuer une analyse des besoins pour le projet et dresser l'inventaire des ressources disponibles à l'interne (matérielles, humaines et financières)	<input type="checkbox"/>
Respecter les étapes de la chaîne de numérisation	<input type="checkbox"/>

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	<input type="checkbox"/>
Classer les documents papier et numériques selon le même plan de classification	<input type="checkbox"/>
Identifier les documents de façon détaillée et significative	<input type="checkbox"/>
Ajouter des métadonnées aux documents numériques dès leur création	<input type="checkbox"/>
Adopter des méthodes de destruction adéquates qui assurent la protection des renseignements personnels et confidentiels <ul style="list-style-type: none">• Déchiquetage à privilégier	<input type="checkbox"/>

Exigences techniques – Recommandations

S'assurer d'avoir et d'utiliser le matériel et les outils adéquats	<input type="checkbox"/>
Choisir le support en fonction de la quantité de documents, la rapidité d'accès aux documents, les coûts et les besoins de sécurité	<input type="checkbox"/>
Choisir une résolution et un mode d'image favorisant la conservation permanente des documents numériques	<input type="checkbox"/>
Choisir des formats sans compression, ou avec une compression sans perte et non soumis à un brevet, favorisant la conservation permanente des documents numériques	<input type="checkbox"/>

Bibliographie

De la théorie à la pratique – Didacticiel d'imagerie numérique, Cornell University,
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., 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/A : l'essentiel – archivage numérique à long terme avec PDF, par Olaf Drümmer, Alexandra Oettler et Dietrich von Seggern, Berlin, PDF/A Competence Center, 2007, 72 p.

Recommandations relatives à la gravure, à la conservation et à l'évaluation des CD-R, Paris, Direction des archives de France, mars 2005, 22 p.