

La numérisation des documents

Méthodes et recommandations

Version revue et corrigée

Mai 2012

*Bibliothèque
et Archives
nationales*

Québec 

Table des matières

Introduction	3
1. Exigences archivistiques et cadre juridique	5
1.2. Intégrité et valeur juridique.....	6
1.3. Documentation du processus de numérisation	7
1.3.1. Déclaration de numérisation	8
1.3.2. Procédure de numérisation.....	9
1.4. Destruction des documents sources	10
1.4.1. Généralités.....	10
1.4.2. Destruction des documents inactifs ayant fait l'objet d'une numérisation	10
1.5. Respect du droit d'auteur	11
1.6. Accès à l'information	12
1.7. Protection des renseignements personnels et confidentiels.....	13
2. Projet de numérisation	14
2.1. Définition de projet : les objectifs visés.....	14
2.2. Choix des documents à numériser et échéancier du projet.....	14
2.3. Analyse des besoins et des ressources disponibles.....	15
2.4. Numérisation des documents.....	16
2.4.1. Préparation matérielle des documents	17
2.4.2. Réalisation de projets pilotes	17
2.4.3. Numérisation des documents	17
2.4.4. Contrôle de la quantité et de la qualité	17
3. Exigences techniques	20
3.1. Outils de numérisation.....	20
3.1.1. Ordinateur	20
3.1.2. Écran à haute résolution et carte vidéo	21
3.1.3. Numériseur.....	21
3.1.4. Logiciel de numérisation et de reconnaissance optique de caractères.....	23
3.2. Résolution et mode de l'image	23
3.3. Formats de conservation des fichiers et compression.....	24
3.3.1. TIFF non compressé.....	24
3.3.2. PDF/A.....	25
3.3.3. XML.....	26
3.3.4. Compression.....	26
3.4. Supports de stockage et protection des fichiers	26
3.4.1. Supports optiques	26
3.4.2. Supports magnétiques	28
3.4.3. Choix de formats et de supports de conservation	28
4. Recommandations pour la numérisation de différents supports.....	29
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 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 guide se veut un outil pratique. 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 organismes publics du Québec. Il aborde l'activité de numérisation sous l'angle de la gestion de projet et de la gestion documentaire. Il est destiné en priorité au personnel des organismes publics assujettis à la Loi sur les archives.

Ce guide a pour principaux objectifs :

- d'informer les organismes en ce qui concerne les exigences archivistiques et les obligations juridiques entourant tout projet de numérisation;
- de fournir des recommandations techniques pour la numérisation de différents types de documents sur divers supports;
- de favoriser la conservation des documents numériques et, dans certains cas et à certaines conditions, l'élimination des documents sources, et ce, peu importe leur stade de vie (documents actifs, semi-actifs et inactifs à conservation permanente).

Ce guide compte quatre sections. La première fait état des exigences archivistiques et des obligations légales guidant toute opération de numérisation. Une attention particulière est portée aux impacts 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 deuxième section du guide présente les différentes étapes d'un projet de numérisation, de la conception à la réalisation. La troisième section expose les différents aspects techniques associés à tout projet de numérisation. On y aborde entre autres les outils de numérisation, la résolution et le mode de l'image, les formats de conservation ainsi que les supports de stockage et de protection des fichiers. Quant à la quatrième et dernière section du guide, elle fait état des recommandations techniques proposées par BAnQ concernant la numérisation de différents types de documents. Ces recommandations y sont présentées sous forme de tableaux.

Le guide comporte aussi en annexe un aide-mémoire qui présente une synthèse des exigences archivistiques, des obligations juridiques ainsi que des principales recommandations de BAnQ applicables à tout projet de numérisation et à la destruction des documents sources. Les organismes peuvent utiliser cet aide-mémoire comme 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.

La Direction générale des archives de BAnQ tient à remercier Hélène Simard, du Service de la gestion intégrée des documents (SGID) du Centre de services partagés du Québec (CSPQ), et Julie Bouchard, du Centre de numérisation ministériel du ministère des Ressources naturelles et de la Faune (MRNF), pour leur contribution à la refonte de ce guide. Nous espérons qu'il facilitera la tâche des organismes désirant mettre sur pied des projets de numérisation.

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.

1. Exigences archivistiques et cadre juridique

Pour lancer une activité de numérisation, il faut connaître les exigences archivistiques et les obligations juridiques qui s’y rattachent, et il faut les respecter. Dans cette partie, nous présentons les principales composantes dont il faut tenir compte.

1.1. Modification au calendrier de conservation

La numérisation des documents actifs et semi-actifs entraîne inévitablement la révision des règles de conservation des séries documentaires touchées par cette opération dans le but :

- d’inscrire le nouveau support au calendrier de conservation et de 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 d’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 sources et, préférablement, avant le début du projet de numérisation¹. C’est uniquement par cette approbation qu’il devient possible de détruire les documents sources actifs et semi-actifs visés par une opération de numérisation. En ce qui concerne la destruction des documents inactifs à conservation permanente qui ont été numérisés, des modalités particulières sont prévues dans la Loi sur les archives. Pour plus d’information à ce sujet, veuillez consulter la section 1.4.2 – *Destruction des documents inactifs ayant fait l’objet d’une 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 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. Pour obtenir de plus amples renseignements sur la soumission et sur l’approbation des calendriers de conservation à BAnQ, veuillez consulter la section « Accueil / Espace professionnel / Archiviste / Services aux organismes publics » sur le portail de BAnQ, à l’adresse suivante : bang.qc.ca/services/archivistique_qed/public/index.html.

Transfert de support :

DÉLAI 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 vérification exhaustive du contrôle de la qualité.									

1.2. Intégrité et valeur juridique²

Un document possède une valeur juridique s'il est susceptible :

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

Ainsi, pour qu'un document numérisé ait la même valeur juridique que le document source analogique, l'intégrité de celui-ci doit être assurée.

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

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

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. La LCCJTI précise que, lors du transfert d'information d'un support vers un autre, l'intégrité du document est maintenue, même en présence de différences touchant :

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

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

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).

stabilité et de pérennité adéquate suivant les normes en vigueur. Par ailleurs, le choix d'un support ou d'une technologie pour consigner de l'information est libre. Par conséquent, le support et la technologie utilisée n'ont pas d'incidence sur la valeur juridique du document.

Lorsque l'intégrité est assurée, 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. 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 original et le remplacer³. Ainsi, il est important de mettre en œuvre des mesures de sécurité afin de maintenir l'intégrité du document numérisé pendant tout son cycle de vie.

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 par la LCCJTI 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.

La documentation doit être conservée durant tout le cycle de vie du document numérisé, car elle constitue la seule preuve attestant des mesures prises pour assurer l'intégrité des documents. Il s'avère donc essentiel d'assurer le lien entre les documents numérisés et la documentation du processus de numérisation.

La documentation du processus de numérisation peut être jointe au document résultant du transfert à ses éléments structurants ou à son support. Elle peut prendre plusieurs formes selon le contexte, la fréquence et les objectifs poursuivis par la numérisation.

La documentation peut, par exemple, prendre la forme 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 par 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 multiples utilisateurs ou pour un organisme qui se dote d'un service permanent de numérisation. Ces

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).

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). Elle s'avère toutefois peu pratique lorsqu'un nombre important d'employés d'une organisation sont autorisés à numériser des documents dans le cadre de leurs tâches quotidiennes. Par ailleurs, soulignons que le principal inconvénient d'une déclaration de numérisation réside dans le fait qu'il est parfois difficile de maintenir le lien entre cette dernière et les documents numérisés, et ce, à moyen et à long terme.

Lorsqu'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 suivants⁵ :

- 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 (au besoin);
- 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.

5. Les organismes publics peuvent s'inspirer de ces éléments dans le but de réaliser leur propre déclaration de numérisation.

1.3.2. Procédure de numérisation

Une procédure de numérisation institutionnelle représente un excellent moyen de documenter le processus de numérisation puisqu'elle permet de définir les tâches, les obligations juridiques et les exigences techniques⁶. Elle peut être utilisée quand la numérisation est effectuée à la pièce par le personnel de l'organisme ou lorsque celui-ci possède une équipe et une infrastructure consacrées à la numérisation. Une procédure de numérisation permet en fait d'alléger l'ensemble du processus de numérisation.

Une procédure de numérisation doit minimalement comporter les éléments suivants :

- la portée et le but de la procédure;
- le nom des personnes ou des groupes de personnes habilités à numériser des documents;
- les responsabilités des différents intervenants;
- les mesures à prendre en vue de la préparation physique des dossiers et des documents avant le début d'une opération de numérisation;
- les outils technologiques et le procédé de transfert utilisés (sécurité, type, caractéristiques et configuration du numériseur, logiciels employés et garanties offertes par ces derniers en matière d'intégrité);
- les choix de traitement, les résolutions, les formats de numérisation et la documentation afférente à produire;
- les politiques et les 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.

Toutefois, 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 ou d'un registre de numérisation (exemples : papier ou fichier tableur .xls ou .scv) dans lequel on trouve le nom de la personne qui a réalisé la numérisation, la date de ce travail, la résolution de la numérisation, les équipements utilisés ainsi que la liste des documents numérisés. Elle peut également être complétée par l'ajout direct de métadonnées au document numérisé.

6. Qui fait quoi, comment, avec quels outils et selon quelles normes.

1.4. Destruction des documents sources⁷

1.4.1. Généralités

À la suite de la numérisation des documents sources, il est possible de les détruire ou de les conserver. Toutefois, si vous décidez de procéder à la destruction de ces documents, il faut au préalable avoir contrôlé l'intégrité des documents, avoir documenté et validé l'opération de transfert, respecté le calendrier de conservation⁸ et protégé les renseignements personnels ou confidentiels lors de l'opération.

Afin de s'assurer que la destruction des documents sources soit faite de façon sécuritaire, 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. De plus, un registre indiquant la date et le procédé de destruction utilisé, le nom de la personne ou du tiers responsable de la destruction, le lieu de la destruction et le nom de la personne responsable du transport des documents pourrait être créé.

Pour obtenir de plus amples renseignements concernant la destruction des documents, nous vous invitons à consulter les documents suivants :

- le *Guide de destruction sécuritaire de l'information*, produit par le ministère des Services gouvernementaux en 2009;
- la norme ISO 15489-2001 – Information et documentation – *Records Management*.

Toutefois, certains documents sources peuvent devoir être conservés après leur numérisation puisque leur support d'origine possède une valeur archivistique, historique ou patrimoniale. Ces valeurs sont établies selon des critères déterminés par règlement du gouvernement⁹. 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 et de la diffusion.

1.4.2. Destruction des documents inactifs ayant fait l'objet d'une numérisation

La Loi sur les archives prévoit que tout « *organisme public doit établir et tenir à jour un calendrier de conservation qui détermine les périodes d'utilisation et les supports de conservation de ses documents actifs et semi-actifs et qui indique quels documents inactifs sont conservés de manière permanente et lesquels*

7. Loi sur les archives (L.R.Q., c. A-21.1, art. 18) et Loi concernant le cadre juridique des technologies de l'information (L.R.Q, c. C-1.1, art. 20 et 69).

8. Le respect du calendrier de conservation s'applique uniquement aux documents actifs et semi-actifs.

9. À ce jour, ces critères n'ont pas été énoncés par le gouvernement.

sont éliminés¹⁰ ».

Il faut donc comprendre que le calendrier de conservation ne fait que désigner les documents qui doivent être détruits et ceux qui doivent être conservés de manière permanente. Une fois le stade semi-actif terminé, les documents à conservation permanente **ne sont plus gérés par le calendrier de conservation** mais plutôt soumis à la portée des articles 14 à 20 de la Loi sur les archives.

En d'autres termes, le fait d'autoriser, au moyen du calendrier de conservation, la destruction de documents actifs ou semi-actifs qui ont été numérisés **ne signifie pas que cette autorisation peut s'appliquer aux documents de cette même série qui ont déjà atteint le stade inactif**. Pour pouvoir détruire des documents inactifs qui ont été numérisés, un organisme public doit en demander l'autorisation à BAnQ en vertu de l'article 18 de la Loi sur les archives, qui précise clairement ceci : « *Nul ne peut aliéner, éliminer ou modifier des documents inactifs d'un organisme public destinés à être conservés de manière permanente. Toutefois, Bibliothèque et Archives nationales peut autoriser l'élimination de tels documents s'ils ont été reproduits sur un autre support ou si elle estime qu'ils sont irrémédiablement détériorés ou qu'il n'est plus utile de les conserver.* »

Pour plus d'information à ce sujet, veuillez consulter le document intitulé *Numérisation de substitution de documents inactifs à conservation permanente – Aide-mémoire pour l'application de l'article 18 de la Loi sur les archives* (L.R.Q., chapitre A-21.1).

1.5. Respect du droit d'auteur¹¹

Qu'il s'agisse de documents actifs, semi-actifs et inactifs à conservation permanente, tout organisme qui désire en numériser doit s'assurer de détenir les droits d'auteur sur les documents à reproduire ou d'avoir recueilli le consentement du titulaire des droits puisque toute reproduction est un droit exclusif de l'auteur.

Dans le contexte de la gestion de documents administratifs, il est **d'usage courant** de reproduire les documents 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.

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

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

1.6. Accès à l'information¹²

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.

Pour ce faire, l'utilisation d'un système de gestion intégrée des documents est la voie à privilégier. En l'absence d'un tel système, le plan de classification peut être reproduit sur un disque 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 pour les documents, en plus de rendre possible l'application des mêmes règles d'accès aux documents.

Pour assurer le repérage des documents, il est nécessaire de les identifier de 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, de même que le maintien de leur intégrité. Ces ajouts peuvent se faire directement dans le fichier ou sur une fiche de métadonnées liée au dossier ou au document¹³.

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 la nature et l'ordre de présentation des éléments d'information. 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.

12. 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, 16.1, 18 à 41, 54, 55, 56, 57, 59 et 63.1).

13. Pour obtenir de plus amples renseignements sur les métadonnées à ajouter aux dossiers et aux documents, 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.

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 sur les supports de conservation sont présentées dans la section 3 – *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 aux répertoires, de façon à ce qu'une personne qui, par exemple, ne possède pas les droits d'accès au document papier ne puisse pas davantage avoir accès à la version numérisée.

1.7. Protection des renseignements personnels et confidentiels¹⁴

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

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

Plusieurs articles de loi précisent ce que sont et ce que ne sont pas des 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.
- les avis juridiques et les recommandations;
- les versions préliminaires et les projets de textes législatifs;
- les analyses, les mémoires et les délibérations;

14. 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).

- les documents d'un ancien ministre versés dans son fonds d'archives privées et confiés au service d'archives d'un organisme public.

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 d'accès physique 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 la destruction sécuritaire des documents, veuillez consulter la section 1.4 – *Destruction des documents sources*.

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 visés

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 ce, dans un format qui assure sa conservation permanente, entraîne des frais supplémentaires. Toutefois, 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 les documents ou les séries documentaires qui seront numérisés, l'ordre dans lequel ce travail sera fait ainsi que la durée du 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 ou de conservation à long terme, 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 stade 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 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 (numériseurs, ordinateurs, logiciels, réseautique et serveurs, locaux et ameublement adaptés, matériel), 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 lieux 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 de l'équipement et des outils qui sont en constante évolution.

Il faut également porter une attention particulière aux normes et aux formats de numérisation ainsi qu'aux supports sur lesquels seront conservés les documents numérisés. Par exemple, pour les documents destinés à une conservation permanente, le recours à des formats ouverts, non compressés, à une résolution élevée ainsi qu'à des supports reconnus pour la conservation à long terme peut 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 sur les supports de conservation, veuillez consulter les sections 3.3 – *Formats de conservation des fichiers et compression* et 3.4 – *Supports de stockage et protection des fichiers*.

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.

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 généralement 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 création d'une expertise institutionnelle. Le recours à l'expertise d'une firme externe présente lui aussi des avantages, notamment l'accès à du personnel possédant une plus grande expérience et à du matériel spécialisé, la rapidité d'exécution, un coût fixe par page ou par document numérisé, 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. Au final, que l'opération soit effectuée à l'interne ou à l'externe, l'important est de s'assurer que le procédé de numérisation soit effectué en conformité avec la Loi sur les archives et la Loi concernant le cadre juridique des technologies de l'information. Pour plus de détails sur cette dernière loi, veuillez consulter sa version commentée à l'adresse indiquée ci-dessous :

msg.gouv.qc.ca/gel/loi_ti/index.html

Pour tout projet de numérisation, il faut respecter les étapes suivantes :

- Préparation matérielle des documents avant et après la numérisation
- Réalisation de projets pilotes respectant les obligations juridiques et, au besoin, apport de correctifs
- Numérisation des documents
- Vérification de la qualité, tant en cours de projet qu'une fois celui-ci terminé
- Destruction ou conservation, selon les cas, des documents sources

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 de l'équipement. Il faut classer les pages ou les documents originaux dans le bon ordre, les déplier au besoin et retirer tout élément pouvant nuire à l'opération (reliures à anneaux, agrafes, trombones, papillons autocollants, éléments d'encadrement, etc.). C'est également à ce moment que peuvent être comptées les pages ou la quantité de documents originaux à numériser.

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

2.4.2. Réalisation de projets pilotes

Avant d'entreprendre la numérisation, il est indispensable de réaliser des projets pilotes afin de tester le processus de numérisation, l'équipement, le logiciel, le choix de traitement, la résolution et le format de conservation des images, les méthodes de contrôle des résultats, les mécanismes de normalisation, d'ajout de métadonnées et de création de copies de sûreté, 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 en prenant soin de manipuler les documents sources 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 sources, 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 tous les documents ou toutes les pages d'un document ont été numérisés. Pour ce faire, il est possible de recourir à des techniques manuelles ou automatisées. La méthode traditionnelle consiste à compter manuellement les pages ou les documents avant la numérisation, puis à comparer le résultat obtenu avec le nombre de pages ou de documents effectivement numérisés. Certains numériseurs et logiciels de numérisation offrent des possibilités automatiques de comptage. 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. De plus, il est important de s'assurer que le recours à des méthodes automatiques de comptage ne laissent pas de traces sur les documents sources.

Contrôle de la qualité des documents numérisés

Lors de la numérisation des documents, il est essentiel de contrôler l'intégrité de leur contenu. Il faut s'assurer que le texte et les annotations sont bien lisibles et visibles, qu'aucune phrase, image, illustration ou ligne n'est coupée, que la résolution est suffisamment élevée, etc. Bref, il est important 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, ainsi que 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, notamment celles-ci :

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

Dans les cas de travaux de numérisation associés à une destruction de documents sources, 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 dans le contrat d'exécution.

Numérisation et reconnaissance optique de caractères

Sauf dans le cas des manuscrits (compte tenu de ce que permet la technologie à l'heure actuelle), il est possible d'appliquer à toute image numérisée d'un document contenant du texte un procédé de reconnaissance optique de caractères (OCR), et ce, peu importe le format de sauvegarde du fichier image. Cette reconnaissance des formes peut être réalisée pendant ou après la numérisation. 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. Au besoin, le résultat peut être sauvegardé dans un fichier en mode texte (.txt, .doc, .rtf, etc.), ce qui permet d'indexer le contenu du document et d'effectuer de la reconnaissance plein texte, de la capture des données et, s'il y a lieu, des modifications.

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 des formes n'est évidemment pas parfait. Enfin, il faut sauvegarder les documents dans un format reconnu pour la conservation permanente (voir les sections 3.3 – *Formats de conservation des fichiers et compression* et 4 – *Recommandations pour la numérisation de différents supports*).

Notez 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 on souhaite obtenir des résultats de recherche fiables, on doit effectuer une vérification exhaustive de la reconnaissance. La lecture optique des textes et le contrôle serré des résultats peuvent entraîner des frais considérables. Il est donc important de bien évaluer au préalable la nécessité d'une telle opération.

3. 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 ainsi que les supports de stockage et de protection des fichiers.

3.1. 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 se conformer aux exigences suivantes :

- s'assurer que les outils utilisés sont compatibles avec les diverses constituantes du processus;
- choisir du matériel standard de qualité, largement répandu et répondant aux normes et aux objectifs établis. Si la numérisation est effectuée par un fournisseur, celui-ci doit offrir du 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 mise à jour et de conversion des fichiers et apporter des améliorations à intervalles réguliers afin que l'information numérisée soit toujours accessible;
- s'assurer que le personnel soit bien formé et mettre sur pied un programme de formation;
- s'assurer du degré de sûreté du système et des données.

3.1.1. 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 les logiciels utiles au projet et connaître leurs caractéristiques. Les fabricants de logiciels et de périphériques indiquent les caractéristiques minimales dont un ordinateur doit être doté pour faire fonctionner leurs produits. Il est généralement recommandé 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 caractéristiques des logiciels et des périphériques qu'on veut acquérir. Il faut tenir compte des éléments suivants :

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

3.1.2. Écran à haute résolution et carte vidéo

L'écran est une composante essentielle à la vérification du résultat de la numérisation de textes et d'images. Afin de réduire la fatigue oculaire du personnel chargé de la numérisation, il est important d'opter pour un moniteur de bonnes dimensions. 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 couleurs, il peut s'avérer utile de se procurer une carte graphique de meilleure qualité que celle de série qui est vendue avec l'ordinateur.

3.1.3. Numériseur

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

Il faut aussi prendre en considération :

- la résolution optique en points par pouce (PPP ou DPI en anglais);
- le spectre du numériseur, qui doit pouvoir rendre toutes les couleurs du document original;

- le traitement des petits formats, des formats hors norme, des formats continus, des grands formats, des copies carbone, des papiers épais, etc.;
- le comptage automatique des pages;
- le cycle de vie de l'appareil et le soutien technique dont il a besoin.

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 reproduire du texte et des images non transparentes ainsi que des objets transparents comme les diapositives 35 mm et les négatifs. Si l'on a un grand nombre de diapositives ou de négatifs 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 pendant 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 microfiches, les documents reliés et les grands formats, par exemple les cartes et les plans. À moins que le volume des documents à numériser soit très important, il est

généralement préférable de confier la numérisation à un fournisseur externe, car l'emploi de ces appareils exige un savoir-faire particulier.

3.1.4. 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 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 son degré de perfectionnement technologique;
- sa convivialité;
- ses possibilités d'adaptation à des besoins spécifiques;
- les formats offerts pour l'importation et pour 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 un procédé de reconnaissance optique des caractères (OCR) aux documents textuels (sauf les manuscrits). Si le volume de 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 pas pu reconnaître, en plus de pouvoir traiter les textes disposés en plusieurs colonnes.

3.2. 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 en anglais). 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 de plus à la qualité perceptible des images et constitue un gaspillage de temps (pour la numérisation) et d'espace disque (pour le stockage des fichiers).

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

Les modes colorimétriques :

- 8 bits couleurs ou tons de gris
- 24 bits vraies couleurs
- 1 bit noir et blanc (bitonal)

Le mode 8 bits couleurs ou tons de gris offre 256 tons de gris ou couleurs différentes. Il est recommandé pour la plupart des documents monochromes, textuels, graphiques ou photographiques. Notez que le mode 8 bits couleurs est généralement suffisant pour les documents comportant quelques éléments en couleurs (en-têtes, signatures, lignes, dessins, etc.).

Le mode 24 bits vraies couleurs offre 16 millions de couleurs et s'impose pour tout ce qui nécessite un rendu de très haute qualité en couleurs.

Quant au mode 1 bit noir et blanc, il restitue approximativement l'original. Il n'est donc pas toujours recommandé, même s'il est le plus avantageux du point de vue de la taille des fichiers. L'utilisation du mode bitonal peut toutefois convenir dans les cas suivants :

- lorsque l'on numérise des documents en excellent état;
- lorsque l'on prévoit utiliser des logiciels spécialisés (exemple : Photoshop) pour rehausser le rendu de l'image.

3.3. Formats de conservation des fichiers et compression

Il existe un grand nombre de formats de conservation de fichiers, mais très peu d'entre eux sont recommandés pour la conservation à long terme.

BAnQ recommande présentement les trois formats suivants pour la conservation à long terme de documents numérisés : TIFF non compressé, PDF/A et XML.

3.3.1. 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 d'images numériques.

En général, il n'est pas recommandé d'utiliser des formats propriétaires comme formats de conservation¹⁵. Toutefois, en raison de sa popularité dans le domaine de l'imagerie numérique et au sein d'organisations vouées à la sauvegarde du patrimoine, le format TIFF demeure malgré tout un format recommandable pour la conservation à long terme d'images numériques.

3.3.2. PDF/A

Le format PDF/A est défini par la norme ISO 19005. Il s'agit d'un format adapté pour la conservation permanente et la diffusion de documents. Il préserve la mise en page, les polices de caractères et la mise en forme. De plus, il s'agit d'un standard ouvert et libre de droits. La raison d'être du PDF/A est d'assurer la conservation à long terme de documents.

Il existe deux versions du format PDF/A, le PDF/A-1 et le PDF/A-2. Le format PDF/A-1 comporte deux niveaux de conformité, le PDF/A-1a et le PDF/A-1b. Le niveau 1a préserve la conformité de la structure logique du document. Il représente la forme la plus complète de la norme ISO 19005-1 (conformité intégrale). Le niveau 1b se limite à la conformité visuelle du document. Ainsi, pour les documents qu'on veut uniquement numériser en mode image, le niveau 1b est suffisant. Il en va de même pour les documents numérisés en mode image dont une lecture optique (OCR) est réalisée par la suite. Pour la conservation de documents structurés (par exemple, les fichiers produits par les logiciels de bureautique comme Word, Excel, PowerPoint, Open Office, etc.), il est recommandé d'utiliser le PDF/A-1a. Si on utilise les versions les plus récentes de ces logiciels, on peut aussi envisager de les sauvegarder en format XML. À ce sujet, veuillez consulter la section 3.3.3 – XML.

Le format PDF/A-2 ne remplace pas le PDF/A-1. Il permet toutefois d'utiliser certaines fonctions des dernières versions du format PDF qui ne sont pas autorisées sous le format PDF/A-1. Le format PDF/A-2 permet, par exemple, l'utilisation du format JPEG 2000, des effets de transparence et des couches d'annotations, l'empaquetage des polices OpenType ainsi que le regroupement de plusieurs documents en format PDF/A dans un seul et même document (porte-documents). Contrairement au PDF/A-1, le PDF/A-2 comprend un troisième niveau de conformité, le PDF/A-2 u. Outre la préservation de la conformité de l'aspect visuel du document, ce niveau offre la possibilité de récupérer le texte en Unicode.

Comme le PDF/A-2 ne remplace pas le PDF/A-1, tous les documents convertis en PDF/A-1 demeurent acceptables pour la conservation permanente de documents. Il n'y a aucune raison, outre celle de profiter des options supplémentaires qu'offre la version 2, de convertir en PDF/A-2 des documents

15. Les formats propriétaires sont des types de fichiers dont les spécifications sont contrôlées par des entités privées. L'utilisation de ces fichiers comme formats de conservation peut poser un certain risque puisque les spécifications de ces formats ne sont pas rendues publiques.

déjà numérisés en PDF/A-1, pas plus que de créer des documents en PDF/A-2 plutôt qu'en PDF/A-1 si aucun des avantages du PDF/A-2 n'est utilisé.

3.3.3. 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 d'utilisation des métadonnées nécessaires à la gestion de 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.

L'utilisation du format XML ou de formats basés sur l'utilisation de schémas XML, par exemple Open Document Format (.odf) et OpenXML (.docx), est particulièrement appropriée pour créer et conserver des documents de type traitement de texte nés numériques ou qui résultent d'une reconnaissance optique de caractères (OCR).

3.3.4. Compression

Pour réduire la taille du résultat de la numérisation, il existe des formats de fichiers compressés. La plupart d'entre eux (exemples : JPEG ou TIFF avec compression G4) entraînent une dégradation des fichiers originaux. D'autres formats, comme le JPEG 2000 et le TIFF avec compression LZW, n'entraînent aucune perte. Dans le cadre de projets de numérisation de documents, il est recommandé d'utiliser des formats non compressés pour les fichiers maîtres afin de faciliter leur conservation à long terme. Si le recours à des formats compressés est nécessaire, seuls les formats n'entraînant aucune perte sont acceptables. Les formats compressés avec perte ne doivent être utilisés que dans les cas de copies de diffusion.

3.4. Supports de stockage et protection des fichiers

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

3.4.1. Supports optiques

Les supports optiques les plus connus sont les CD-ROM (Compact Disc ou disques optiques compacts), les DVD (Digital Versatile Disc) et les BD (Blu-ray 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). Pour des raisons d'instabilité, il vaut mieux

éviter les DVD double couche et les disques réinscriptibles DVD-RW (DVD Rewritable). Quant aux disques Blu-ray, ils sont déconseillés comme supports de conservation en raison de la fragilité de la fine couche de substrat transparent protégeant le film métallique sur lequel l'information est gravée. Ainsi, il est préférable de choisir des disques inscriptibles (DVD-R : DVD Recordable) de qualité « Archives » (exemples : Verbatim Ultralife Gold ou Mitsui / MAM-A Gold). Ces disques sont plus coûteux que les DVD-R 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 celles-ci ne se retrouvent pas à la périphérie du disque, une zone plus fragile en raison de son exposition aux traces de doigts;
- graver les fichiers à la vitesse recommandée par le fabricant de disques puisqu'une vitesse supérieure de gravure peut entraîner la perte de données;
- avant d'entreprendre la gravure, fermer tous les programmes inutiles et désactiver toutes les autres applications sur l'ordinateur;
- conserver les disques dans 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 (exemple : tous les trois à cinq ans);
- planifier la migration des données sur un autre support (repiquage) à tous les cinq ans.

3.4.2. 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 dans le cas des 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.). (DLT = Digital Linear Tape). Elles sont peu pratiques pour l'accès courant aux données, car seule une lecture séquentielle des données est possible. Cette caractéristique peut rendre la recherche de documents plus lente et ardue. Ces cartouches offrent toutefois 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.

3.4.3. 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 tenir compte du fait qu'à long terme, des transferts et des migrations de formats et de supports s'imposeront nécessairement pour assurer l'intégrité et la permanence de l'accès aux données.

4. Recommandations pour la numérisation de différents supports

Les tableaux ci-dessous résument les choix de traitement, les résolutions d'usage et les formats de conservation pour différents types de documents. Retenez que l'objectif qui sous-tend ces recommandations est d'offrir aux organismes publics des balises visant à assurer la création de documents numérisés de qualité dont l'intégrité peut être assurée. Bien que considérée comme un minimum par BAnQ pour atteindre cet objectif, l'application de ces balises peut varier en fonction de la qualité et des caractéristiques physiques des documents et de l'équipement utilisé.

Rappelons qu'une résolution de 300 PPP est considérée comme un minimum suffisant pour reproduire un document à sa taille originale.

Support d'origine	Choix de traitement	Résolution	Formats de conservation
Documents textuels sur papier	Recommandés : <ul style="list-style-type: none"> • 8 bits tons de gris • 8 bits couleurs 	<ul style="list-style-type: none"> • 300 PPP 	Recommandés : <ul style="list-style-type: none"> • TIFF non compressé • PDF/A • XML
Documents textuels sur microfilms et sur microfiches	Recommandé : <ul style="list-style-type: none"> • 8 bits tons de gris 	<ul style="list-style-type: none"> • Le taux varie en fonction du facteur de réduction utilisé (exemples : 18 x, 24 x). 	Recommandés : <ul style="list-style-type: none"> • TIFF non compressé • PDF/A
Cartes et plans	Recommandés : <ul style="list-style-type: none"> • 8 bits tons de gris pour les documents en noir et blanc • 24 bits couleurs pour les documents en couleurs 	<ul style="list-style-type: none"> • 300 PPP pour les formats plus grands que 11 x 17 • 600 PPP pour les formats plus petits que 11 x 17 ou pour le géoréférencement des plans 	Recommandés : <ul style="list-style-type: none"> • TIFF non compressé • GeoTIFF pour les plans géoréférencés

Support d'origine	Choix de traitement	Résolution	Formats de conservation
Épreuves photographiques	Recommandés : <ul style="list-style-type: none"> • 8 bits tons de gris pour les épreuves en noir et blanc • 24 bits couleurs pour les épreuves couleurs, incluant les épreuves sépia 	<ul style="list-style-type: none"> • 300 PPP pour les formats 8 x 10 et les plus grands formats • 600 PPP pour les formats plus petits que 8 x 10 	Recommandé : <ul style="list-style-type: none"> • TIFF non compressé
Négatifs	Recommandés : <ul style="list-style-type: none"> • 8 bits tons de gris • 24 bits couleurs 	<ul style="list-style-type: none"> • 2400 PPP pour du 35 mm • 1200 PPP pour le format 2 1/2 x 2 1/2 • 600 PPP pour les formats 4 x 5 et 5 x 7 • 300 PPP pour le format 8 x 10 	Recommandé : <ul style="list-style-type: none"> • TIFF non compressé
Diapositives	Recommandés : <ul style="list-style-type: none"> • 8 bits tons de gris pour les diapositives en noir et blanc • 24 bits couleurs pour les diapositives en couleurs 	<ul style="list-style-type: none"> • 2400 PPP pour le format 35 mm • 1200 PPP pour le format 2 1/4 x 2 1/4 	Recommandé : <ul style="list-style-type: none"> • TIFF non compressé

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 encore, 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 (pour les documents actifs et semi-actifs seulement)* Approbation à obtenir avant la destruction des documents sources, 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 pour les documents actifs et semi-actifs;• avoir obtenu l'autorisation de BAnQ en ce qui concerne la destruction de documents inactifs ayant été numérisés (article 18 de la Loi sur les archives).	<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 ou selon tout autre système de classement 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"/>

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 plan de classification ou selon tout autre système de classement	<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">• Destruction fragmentaire à privilégier (exemple : déchiquetage)	<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 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 en matière 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, préférablement non soumis à un brevet, afin de favoriser la conservation permanente des documents numériques	<input type="checkbox"/>

Bibliographie

ARCHIVES NEW ZEALAND / GOVERNMENT RECORDKEEPING GROUP, *Digitisation Standard*, 2010, archives.govt.nz/sites/default/files/S6_-_Digitisation_Standard_-_PDF_-_Final_-_23_06_08.pdf (consulté le 1^{er} mai 2012).

BÉLANGER, Micheline, Jean Maurice DEMERS et Claude MINOTTO, *Guide d'imagerie numérique – Numérisation de l'information consignée sur des supports traditionnels*, Québec, Archives nationales du Québec, 1998, 67 p.

BIBLIOTHÈQUE ET ARCHIVES CANADA, *Initiatives numériques de BAC*, www.collectionscanada.gc.ca/digital-initiatives/index-f.html (consulté le 1^{er} mai 2012).

BROSSEAU, Kathleen, Mylène CHOQUETTE et Louise RENAUD, *Normes de numérisation de la Société du Musée canadien des civilisations*, version 1.1, mars 2006, www.pro.rcip-chin.gc.ca/sgc-cms/coursenligne-onlinecourses/content/31/MCCnumerisation.pdf (consulté le 1^{er} mai 2012).

CHAUMIER, Jacques, *Document et numérisation – Enjeux techniques, économiques, culturels et sociaux*, Paris, Association des professionnels de l'information et de la documentation, 2006, 119 p.

COMITÉ INTERMINISTÉRIEL SUR LES MÉTADONNÉES, *Profils de métadonnées gouvernementaux – Dossier et documents – Version 2.0*, avril 2009, www.banq.qc.ca/documents/services/archivistique_ged/Profils_metadonnees_gouvernementaux_Dos_Doc_version2_0_avril2009_VF.pdf (consulté le 1^{er} mai 2012).

COMMISSION D'ACCÈS À L'INFORMATION, *Guide pour la destruction des documents renfermant des renseignements personnels*, www.cai.gouv.qc.ca/documents/CAI_FI_destruction.pdf (consulté le 1^{er} mai 2012).

CORNELL UNIVERSITY, *De la théorie à la pratique – Didacticiel d'imagerie numérique*, www.library.cornell.edu/preservation/tutorial-french/contents.html (consulté le 1^{er} mai 2012).

DRÜMMER, Olaf, Alexandra OETTLER et Dietrich VON SEGGERN, *PDF/A : l'essentiel – Archivage numérique à long terme avec PDF*, Berlin, PDF/A Competence Center, 2007, 72 p.

GROUPE DE TRAVAIL MINERVA N^o 6, *Guide des bonnes pratiques – Identification des bonnes pratiques et des centres de compétence – Version 1.3 de mars 2004 élaborée dans le cadre du projet Minerva*, 3 mars 2004, http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bonnespratiques1_3.pdf (consulté le 1^{er} mai 2012).

MAILLET, Dominique, Bibliothèque nationale de France, Département de la conservation, *Éléments techniques pour la numérisation des cartes et plans à la Bibliothèque nationale de France*, www.bnf.fr/documents/MAILLET_jp_num_portulans_2011_presentation.pdf (consulté le 1^{er} mai 2012).

MINISTÈRE DES SERVICES GOUVERNEMENTAUX, *Guide de destruction sécuritaire de l'information*, janvier 2009, www.informa.msss.gouv.qc.ca/Details.aspx?Id=UYsVJSLouly= (consulté le 1^{er} mai 2012).

ORGANISATION INTERNATIONALE DE NORMALISATION, *ISO 19005-1:2005 – 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)*, www.iso.org/iso/fr/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=38920 (consulté le 1^{er} mai 2012).

ORGANISATION INTERNATIONALE DE NORMALISATION, *ISO 19005-2:2011 – Gestion de documents – Format de fichier des documents électroniques pour une conservation à long terme – Partie 2 : Utilisation de l'ISO 32000-1 (PDF/A-2)*, www.iso.org/iso/fr/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50655 (consulté le 1^{er} mai 2012).

ORGANISATION INTERNATIONALE DE NORMALISATION, *ISO 15489-1:2001 – Information et documentation – « Records Management » – Partie 1 : Principes directeurs*, www.iso.org/iso/fr/catalogue_detail.htm?csnumber=31908 (consulté le 1^{er} mai 2012)

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.

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

SANTÉ ET SERVICES SOCIAUX DU QUÉBEC, DIRECTION GÉNÉRALE DE LA COORDINATION, DU FINANCEMENT, DE L'ÉQUIPEMENT ET DES RESSOURCES NATURELLES, *Guide – Catégorisation des documents technologiques en matière de sécurité – Trousse d'outils*, 2009, [msssa4.msss.gouv.qc.ca/extranet/ri.nsf/49dd266bd183416e852566e2005c98b6/8aa8a7d1b1e28b8e852571420056a43f/\\$FILE/MSSS04-021%20Guide%20de%20categorisation%20v1,3%20vapprouvee%202006-10-05.pdf](http://msssa4.msss.gouv.qc.ca/extranet/ri.nsf/49dd266bd183416e852566e2005c98b6/8aa8a7d1b1e28b8e852571420056a43f/$FILE/MSSS04-021%20Guide%20de%20categorisation%20v1,3%20vapprouvee%202006-10-05.pdf) (consulté le 1^{er} mai 2012).

TGE ADONIS, *Le guide des bonnes pratiques numériques – Version 2*, 15 septembre 2011, www.tge-adonis.fr/sites/default/files/ressourcesdoc/guide_des_bonnes_pratiques_v2.pdf (consulté le 1^{er} mai 2012).