Document généré le 20 juin 2025 17:24

Documentation et bibliothèques

La gestion des documents électroniques : le point de vue de la Bibliothèque nationale du Canada

Rolande Blair

Volume 42, numéro 3, juillet–septembre 1996

URI : https://id.erudit.org/iderudit/1033264ar
DOI : https://doi.org/10.7202/1033264ar

Aller au sommaire du numéro

Éditeur(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (imprimé)
2291-8949 (numérique)

Découvrir la revue

Citer cet article

Blair, R. (1996). La gestion des documents électroniques : le point de vue de la Bibliothèque nationale du Canada. *Documentation et bibliothèques*, *42*(3), 127–133. https://doi.org/10.7202/1033264ar

Tous droits réservés © Association pour l'avancement des sciences et des techniques de la documentation (ASTED), 1996

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

https://apropos.erudit.org/fr/usagers/politique-dutilisation/

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

https://www.erudit.org/fr/

La gestion des documents électroniques : le point de vue de la Bibliothèque nationale du Canada

Rolande Blair
Bibliothèque nationale du Canada

De nos jours, les médias font grand état de l'impact des nouvelles technologies sur la société en général et sur l'accès à l'information et le rôle des bibliothèques en particulier. Il ne fait aucun doute que cet impact est majeur. On parle même de révolution dépassant en répercussions sociologiques et économiques celles de la révolution industrielle. Les débuts fracassants de cette «civilisation électronique» sont largement démontrés par l'engouement d'un public de plus en plus large pour la quincaillerie et les produits informatiques de tout type et de tout genre, par la prolifération des industries axées sur la création et l'organisation de l'information et par les fortunes accumulées par les producteurs de logiciels et autres produits reliés à l'informatique.

Ce nouvel environnement «électronique» fait sentir son influence de façon toute particulière dans le secteur de l'édition, celui de la globalisation des marchés et celui de la prolifération des produits et services dits «électroniques». On parle de réalité virtuelle. On s'inquiète et on cherche à assurer la présence canadienne dans ce nouvel environnement. Les gouvernements voient dans ces nouvelles technologies de nouvelles possibilités de se rapprocher du citoyen et d'améliorer la qualité de ses services.

Au coeur de ces nouvelles technologies, il faut noter de façon toute particulière la rapide croissance d'Internet et des technologies qui en favorisent l'accès et l'utilisation. Internet, l'autoroute de l'information, attire l'attention de tous mais plus particulièrement l'intérêt des éducateurs et des bibliothécaires, car elle permet en principe l'accès universel à l'information pourvu que cette information circule sous forme d'octets, sur des infrastructures qui en permettent la transmission rapide et qu'elle soit éventuellement entreposée dans des sites qui en assurent la garde et la gestion.

Cette vision, qui semble simple a priori, repose sur un ensemble de préalables et de facteurs tous plus complexes les uns que les autres. La Bibliothèque nationale du Canada (BNC) a commencé à s'intéresser à cette problématique lors d'une rencontre de planification, tenue en mars 1994, dont le but était de préciser son mandat face à cette émergence de l'électronique. Sa réflexion s'est portée sur deux aspects complémentaires de la question: premièrement, son rôle dans la garde et la gestion des publications électroniques créées par des auteurs et éditeurs canadiens et, deuxièmement, son action comme créateur et diffuseur de documents numérisés. Cet article entend s'étendre davantage sur le premier volet de la question sans négliger tout à fait le deuxième volet qui a une importance stratégique considérable comme on l'a constaté lors du dépôt du rapport du Conseil consultatif sur l'autoroute de l'information.

Le rôle de la BNC dans la garde et la gestion des publications électroniques canadiennes

Le mandat de la BNC compte deux aspects principaux. La BNC joue tout d'abord un rôle primordial dans l'acquisition, la préservation et l'accès au patrimoine publié canadien. La BNC a également un rôle de premier plan dans l'accès à l'information et au patrimoine publié mondial parl'ensemble des citoyens du Canada, quel que soit l'endroit où ils se trouvent, où cette information se trouve ou quel que soit le format sous lequel cette information est disponible.

Le passage actuel à un contexte d'information électronique soulève un ensemble de questions qui modifieront la façon dont la BNC remplira son mandat patrimonial. L'identification, le repérage, l'acquisition, l'accès, le stockage et la conservation des documents électroniques soulèvent des questions qui doivent être traitées et résolues dans un avenir proche. Si on ne s'occupait pas de ces questions, de sérieuses lacunes apparaîtraient dans le patrimoine canadien et les outils de recherche du pays. Vue sous cet angle, la pertinence d'un rôle pour la BNC en ce qui concerne les publications électroniques canadiennes ne fait aucun doute. La BNC se doit d'acquérir et de préserver pour la postérité les publications électroniques canadiennes.

La BNC accepta à priori l'hypothèse que son mandat s'étendait non seulement aux publications sur support matériel comme les disques optiques ou les disquettes par exemple, mais également aux publications purement électroniques, comme celles qui n'existent qu'en octets dans les sites Internet. La nature éphémère de ces publications, appelées à disparaître d'un coup de baguette de leur créateur, soulignait encore davantage, nous semblait-il, l'importance du rôle de sauvegarde de la BNC si cette partie du patrimoine publié canadien se devait d'être préservée. L'accès par tous à toutes les publications canadiennes d'aujourd'hui et d'hier, quel que soit leur format, devenait un défi de taille.

Un projet pilote à la rescousse

Avant d'intégrer le traitement et la gestion des publications électroniques à ses activités régulières, la BNC se devait d'en comprendre tous les aspects. Un projet pilote allait de soi. La BNC dirigea donc, de juin 1994 à juillet 1995, un groupe de travail chargé d'acquérir, d'entreposer et de donner accès à un petit nombre de périodiques électroniques canadiens. Ce projet pilote fut nommé le *Projet Pilote sur les Publications Électroniques* ou plus familièrement le *PPPE* ou *3P-E*.

Les objectifs du projet pilote étaient, entre autres, d'identifier et de comprendre les problèmes auxquels font face les bibliothèques en général et la BNC, en particulier, dans le traitement des collections ordinolingues. Tout au long du projet, la BNC se devait d'accroître son expérience et son expertise dans certaines des technologies et des enjeux techniques mettant en cause l'édition et les publications électroniques, en particulier l'édition électronique dans Internet.

La planification du *3P-E* se fit durant l'été 1994 avec la participation du personnel de plusieurs secteurs de la BNC. Les membres de l'équipe, en plus de se familiariser avec Internet et tous les aspects de la gestion des documents électroniques, ont examiné des aspects connexes tel le problème du droit d'auteur et de la préservation à long terme de ces documents. Douze mois après le début du projet, le groupe de travail soumettait un rapport qui comptait 31 recommandations. Les sections qui suivent feront état des considérations et délibérations les plus importantes survenues en cours d'exercice.

Identification et acquisition des publications électroniques

J'aborderai cette section dans l'ordre où les problèmes sont survenus après que la décision fut prise d'aller de l'avant avec le projet pilote.

Il y a un volume important de documents canadiens de tous types dans Internet. Il fut vite reconnu qu'il était à toutes fins impossible pour quelque bibliothèque que ce soit d'acquérir et de préserver tous ces documents. Le groupe de travail tenta donc, et ce fut là le premier défi auquel il dut faire face, de définir la «publication électronique». On essaya de le faire d'un point de vue essentiellement pratique, cherchant à déterminer le type de publications électroniques que la BNC se devrait d'acquérir et de préserver pour la postérité. Il fut décidé de procéder par la négative, en tentant de déterminer d'emblée certains champs d'exclusion. Furent exclus a priori le courrier électronique, les sites/pages d'accueil WWW ou GOPHER, les bases de données ordinolingues, les archives du FTP (Protocole de transfert de fichiers) et les babillards électroniques. Toutefois, à l'intérieur des sites WWW et GOPHER, il peut y avoir des documents qui ont les caractéristiques de publications traditionnelles, comme le passage par les activités de préparation spécifique associées traditionnellement au processus d'édition. Après moult délibérations, on en arriva, pour le projet, à la définition suivante:

> *Une publication électronique est un document qui résulte de l'acte de publier, où l'information est codée et rendue accessible et intelligible au moyen d'un ordinateur. Il y a deux catégories de publications électroniques: celles qui sont distribuées en copies multiples sur des supports «durs», tels le disque compact, la disquette et la bande magnétique; et celles qui se trouvent dans la mémoire d'un ordinateur central et qui sont accessibles par un réseau de communications. Les publications électroniques font l'objet des mêmes activités de préparation que les publications imprimées traditionnelles.*

On doit comprendre toutefois que le problème de «définition», dans l'éventualité d'une intégration des publications électroniques aux opérations courantes, reste entier. Avec l'émergence de l'électronique, la démarcation entre le document et la publication électronique devient de plus en plus floue. Une définition, à la fois claire et exhaustive, contribuerait à long terme à définir les mandats des divers intervenants qui traiteront l'information électronique.

La seconde étape visa à établir, à l'intérieur des inclusions (des documents répondant à la définition des publications électroniques citée plus haut), un cadre plus restreint qui fixerait les paramètres du projet pilote comme tel. Il s'agissait donc pour les membres de l'équipe d'identifier, pour le projet, le type et le nombre de publications électroniques canadiennes qui permettraient d'acquérir une masse suffisante d'information pour en arriver à certaines conclusions. Il fut décidé que le projet pilote devait cibler de prime abord les «périodiques» électroniques canadiens, plus faciles à identifier et plus appropriés pour un projet pilote parce qu'ils ont une histoire et une continuité. Certains critères de sélection furent appliqués. Les publications d'intérêt purement local furent exclues de même que les publications purement publicitaires. De plus, à cause de contraintes budgétaires, (ce projet disposait d'un budget de fonctionnement de 20 000 $ seulement), le *3P-E* se devait d'exclure toute publication électronique non accessible par un logiciel d'accès gratuit.

Le projet lancé, la chasse aux périodiques électroniques canadiens dans Internet fut ouverte. Le progrès fut lent au début, mais grâce au bouche à oreilles et à la croissance rapide de l'édition électronique dans Internet au Canada, l'intérêt des éditeurs de périodiques électroniques pour ce projet grandit rapidement. Alors même que le nombre visé au début n'était que d'une dizaine de titres, 46 titres étaient enregistrés sur notre serveur à la date du dépôt du rapport final.

Pour la BNC, le simple fait d'acquérir le droit de «pointer» une publication électronique ou d'établir des liens électroniques dynamiques avec une publication stockée sur le site d'un éditeur électronique ne suffit pas à en assurer la conservation. Jusqu'à présent, l'expérience montre que la conservation des publications électroniques ne peut être laissée entre les mains des éditeurs électroniques, tout comme on ne peut laisser aux éditeurs «traditionnels» le soin de conserver des documents sur d'autres supports. La BNC se doit donc d'obtenir un exemplaire électronique des publications électroniques qu'elle a le mandat de garder et de rendre accessible aux chercheurs et lecteurs présents et futurs.

Les dessous de l'acquisition électronique réservèrent des surprises intéressantes. La Direction des acquisitions et services bibliographiques de la BNC fut impliquée dans le repérage des titres répondant aux critères de sélection ainsi que dans la négociation, avec les éditeurs concernés, des droits à l'utilisation de leurs titres aux fins du projet pilote. Ces tâches s'avérèrent très onéreuses en ressources humaines, car il faut ajouter au temps nécessaire pour identifier et localiser les publications le temps requis pour les nombreuses communications avec les éditeurs. On a évalué que deux années-personnes seraient requises pour

le traitement de 500 à 1 000 publications par année.

Durant le projet pilote, le protocole d'acquisition des périodiques fut largement déterminé par le protocole utilisé pour la distribution des périodiques électroniques dans Internet par chacun des éditeurs. Les publications furent donc reçues par courrier électronique, transfert de fichiers ou par des logiciels tels que Mosaic ou Netscape. Les publications HTML furent initialement difficiles à capter à cause du nombre de fichiers associés à un seul document. La création d'un «double/miroir» du document tel qu'il apparaît sur le site hôte fut la solution adoptée. Cependant on recommande, dans ce cas, de prendre périodiquement des instantanés du site hôte et de les entreposer sur le serveur local pour maintenir l'intégrité du titre au cas où le site hôte déciderait d'éliminer certains fichiers/articles/numéros.

Le projet pilote appliqua les procédures traditionnelles de contrôle des périodiques. Cependant, il est évident que la validation du contenu doit être ajoutée aux procédures habituelles dans le cas de documents avec hyperliens pour vérifier que tous les documents associés ont bien été reçus et sont maintenus.

Comme l'expérience a démontré que l'acquisition des publications électroniques s'avère une fonction coûteuse en ressources humaines et comme le nombre et la qualité des publications électroniques augmentent rapidement, la BNC se devra de rechercher, en priorité, des solutions automatisées qui facilitent les procédures d'acquisition et de contrôle des publications.

Un autre sujet auquel le groupe de travail s'attarda assez longuement, vu le mandat spécifique de la BNC à cet égard, fut la question du dépôt légal. Déjà en 1993, les règlements du dépôt légal étaient amendés pour y spécifier l'inclusion des disques compacts et des vidéocassettes publiés au Canada. Devait-on en faire autant pour les publications électroniques? Durant le projet pilote, ledépôt légal ne fut pas invoqué pour l'obtention des titres désirés. Les titres furent plutôt obtenus par le biais d'arrangements avec les éditeurs. La plupart des éditeurs qui participèrent au projet acceptèrent avec plaisir de fournir des copies de leurs publications à la BNC, jugeant que ceci ajoutait à la crédibilité et à l'accessibilité de leurs publications.

Le grand avantage d'étendre le dépôt légal aux publications électroniques est de donner à la BNC l'autorité légale d'acquérir les publications électroniques canadiennes pour la collection nationale. Ce levier ou cette autorité est d'autant plus nécessaire lorsque l'éditeur ou le vendeur impose des frais d'accès à une publication et pourrait être peu enclin à ce que la BNC donne librement accès au document ou le copie pour ses usagers. (Voir la section sur l'accès aux publications électroniques).

Le principal inconvénient de l'extension du dépôt légal aux publications électroniques est que le dépôt légal représente un mode d'acquisition relativement aveugle. Dans notre monde électronique, tout individu avec accès à Internet peut devenir son propre éditeur. À moins que ne soit appliqué un mécanisme quelconque de sélection, le volume de documents et leur complexité submergeront les capacités de la Bibliothèque à faire quoi que ce soit avec ces acquisitions. Le groupe de travail en est vite venu à la conclusion que le recours au dépôt légal devrait se faire de façon sélective et ne s'appliquer qu'aux documents que la BNC veut acquérir pour la collection nationale permanente. Pour l'instant, la Loi de la Bibliothèque nationale du Canada et les Règlements du dépôt légal n'excluent pas les publications électroniques. Toutefois, il y a beaucoup d'incertitude en ce moment concernant le statut de «publications» donné aux documents transmis électroniquement. L'accord de libre échange nord-américain (l'ALENA) a eu pour résultat d'amender la Loi sur le droit d'auteur du Canada (en vigueur le 31 janvier 1994) comme suit:

> *Pour l'application de cette Loi, une «publication» n'inclut pas la communication d'une oeuvre au public par télécommunication...*

L'obtention d'un avis juridique concernant le dépôt légal des publications transmises électroniquement fut conséquemment reconnue comme un préalable au développement d'une stratégie d'acquisition et de gestion de ces documents.

La BNC s'applique présentement à raffiner les critères de sélection qu'elle appliquera aux publications électroniques. En général ces critères tiennent compte des directives générales touchant le développement des collections de la BNC qu'on retrouve dans les *Politiques de gestion des collections de la Bibliothèque nationale du Canada.*

Durant le projet pilote, la BNC s'est limitée à l'acquisition de publications en série transmises sur Internet. Quelques critères généraux furent appliqués lors de la sélection des titres pour le projet pilote. La publication électronique devait être canadienne, c'est-à-dire publiée au Canada et/ou sur un serveur canadien, avec un éditeur canadien ou d'importants contributeurs canadiens.

D'autres critères sont présentement développés. Les monographies y seront évidemment considérées au même titre que les périodiques. L'emphase y sera certainement mise sur la littérature, l'histoire et la musique qui sont les secteurs privilégiés par la BNC. Priorité sera également donnée aux publications gouvernementales, aux titres uniquement disponibles électroniquement et aux publications non commerciales, du moins dans un premier temps. Il va de soi que l'intégration de ce nouveau format aux activités régulières de la BNC se fera sélectivement et par étape. Les ressources de la BNC, dramatiquement réduites ces dernières années, ne permettent aucune autre stratégie.

L'accès aux publications électroniques

Durant le projet pilote, l'accès intellectuel ou bibliographique aux publications électroniques s'est fait à trois niveaux: par le biais de la notice individuelle, de la structure de fichiers et des outils de recherche dans Internet.

La BNC a préparé des notices catalographiques pour tous les titres sélectionnés, incluant de l'information sur l'acquisition électronique du document ainsi que les adresses URL du document,

soit celles du site de la BNC et du site hôte. La zone 856 du MARC et ses sous-zones ont été utilisées à cet effet. L'analyse du sujet des titres, en direct, prit beaucoup de temps à cause de la navette que les catalogueurs ont dû faire entre les fichiers pour différents numéros et articles. Au cours du projet pilote un lien coupé-collé entre l'enregistrement URL d'une notice bibliographique et les serveurs WWW/ GOPHER fut utilisé pour accéder aux documents électroniques. Un lien direct du URL au document sera construit dans l'avenir.

La structure de la banque de documents, les répertoires et les noms de fichiers furent des composantes clés permettant l'accès aux documents entreposés dans notre serveur. Les conventions de désignation des fichiers sont importantes pour l'organisation, le stockage et l'accès à l'information. Les outils de gestion de fichiers, tels ceux utilisés par Windows, DOS ou VMS, existent déjà. Ils furent utilisés dans le cadre de ce projet. Des systèmes de gestion de documents beaucoup plus perfectionnés sont aussi disponibles, mais leur coût dépassait de beaucoup le cadre du *PPPE*.

Le troisième niveau d'accès se fit selon les protocoles Internet, soit le GOPHER et le WWW. Des catégories «sujet» (basées sur Dewey) furent assignées à chaque titre et les titres furent listés par ordre alphabétique sur nos sites WWW et GOPHER. Des index WAIS donnant accès aux titres et au texte de l'ensemble des documents stockés furent également développés. Le projet pilote démontra qu'il serait possible de construire des index plus précis, par exemple d'indexer le plein texte d'un titre particulier ou tous les textes sur un sujet donné. Notre expérience démontra toutefois qu'il serait difficile, sans intervention majeure, de construire des index de tables de matière parce que la plupart des publications électroniques n'ont pas ce niveau de structure de contenu.

Les éditeurs de plusieurs titres électroniques inclus dans le projet pilote avaient déjà demandé un ISSN, avant le démarrage du projet. L'attribution d'un numéro international normalisé des publications en série (ISSN) aux titres se fit donc couramment. Les titres furent conséquemment enregistrés dans la base de données du réseau d'ISSN.

Afin de soutenir et d'aider le travail du catalogueur et de favoriser l'accès aux publicationsélectroniques, il fut recommandé que la BNC analyse et définisse ses exigences concernant l'identification des données. Idéalement, des métadonnées devraient être publiées comme faisant partie intégrante des publications électroniques. Un travail considérable a déjà été effectué dans ce domaine et il faudrait qu'il soit analysé avec et par des éditeurs électroniques et des groupes s'occupant des normes, pour déterminer les exigences précises du Canada et des bibliothèques.

L'accès physique aux collections électroniques s'est avéré plus problématique. La BNC a discuté du sujet durant de nombreuses heures tant en petits groupes que lors de réunions impliquant un plus grand nombre d'employés et de collaborateurs. Fondamentalement la BNC aimerait assurer aux publications électroniques le même niveau d'accès que celui qu'elle assure aux autres publications dans ses collections. Évidemment nos préoccupations sont décuplées par les questions touchant le droit d'auteur, une problématique plus qu'actuelle de l'édition électronique au Canada.

Le groupe de travail étudia toute une gamme de scénarios. Dans tous les scénarios, l'accès au sous-ensemble gratuit de la collection (documents non commerciaux, publications gouvernementales) serait évidemment offert sans frais, que ce soit pour utilisation à l'interne, par le prêt entre bibliothèques ou par Internet. Pour le reste de la collection, il fut envisagé, entre autres, d'imposer des frais à chaque client pour l'accès; de payer les éditeurs pour chaque accès aux documents par les clients de la BNC; de limiter l'accès aux publications payantes à quelques stations interactives installées dans la salle de référence de la BNC; d'offrir l'accès gratuit à quiconque ayant accès à la collection de la BNC par Internet ou par d'autres moyens; et finalement de négocier avec l'éditeur électronique certaines restrictions à l'accès pour une période limitée, période qui serait dictée par la valeur commerciale du document. Tous ces scénarios furent analysés en détail.

Quoique la BNC se doive de réserver ses conclusions jusqu'à ce qu'un avis juridique lui soit donné, il est bon de noter que l'équipe de projet se prononça contre des systèmes ou mesures de contrôle visant à facturer l'usager au profit de l'éditeur ou à comptabiliser l'accès de façon à ce que la BNC puisse payer l'éditeur pour chaque utilisation d'un document particulier. Notre position de base reste ferme. Les publications de la BNC doivent être mises à la disposition des chercheurs ou autres utilisateurs, directement ou par le prêt entre bibliothèques et l'accès à ces documents doit être équitable, abordable et universel.

Dans l'attente d'une position définitive, la BNC doit assurer le suivi des autorisations et des restrictions portant sur l'accès aux titres de sa collection de publications électroniques en enregistrant dans la base de données AMICUS (la base de données de la BNC) le statut des droits d'auteur de chaque publication électronique de sa collection et, s'il s'agit d'une publication gratuite et d'accès libre, les raisons de ce statut.

La gestion, le maintien et la conservation des publications électroniques

Les défis liés à l'acquisition et à la gestion d'une collection de documents électroniques ne s'arrêtent malheureusement pas à son développement, son traitement et son accessibilité. Viennent ajouter à la complexité de l'ensemble toutes les questions touchant le stockage, le maintien et la préservation de la collection.

Tout d'abord, il existe des problèmes liés à **l'entreposage** des collections électroniques. Durant le projet pilote, les publications électroniques furent entreposées et rendues accessibles à partir d'un serveur central géré par les Services de technologie de l'information de la BNC. Cet infoserveur fonctionne selon le système d'exploitation UNIX. On a examiné un certain nombre d'options pour désigner, organiser et stocker les publications électroniques sur le serveur de la BNC. On adopte le gestionnaire de fichiers Microsoft Windows sur les postes de travail et un gestionnaire de fichiers du domaine public (le gratuiciel de la base de données SQL) sur l'infoserveur. Une structure de

stockage des fichiers a été construite, des conventions de nomenclature des fichiers acceptées et des répertoires développés. Les adresses URL furent construites en cours de fonctionnement. À plus long terme, les observations préliminaires indiquent que la solution de conservation devrait être un système de stockage hiérarchique (HSM) : les documents actuels ou couramment demandés sont stockés en direct sur des supports de récupération ultra-rapides; les documents d'accès moins fréquent sont stockés sur des supports presque en direct (le disque optique) et les documents à faible demande sont stockés hors ligne, sur bandes magnétiques par exemple.

Un mot maintenant sur **l'organisation** des documents. Ils devraient être organisés et repérables grâce à une structure de type «base de données» de façon à en assurer l'accès par plusieurs points d'entrée. Le personnel réalisa rapidement que le développement et l'utilisation de répertoires ainsi que de conventions pour la nomenclature des fichiers représentent une tâche importante dont il faut tenir compte dans ce qu'il convient d'appeler la «gestion» des documents électroniques. La gestion de fichiers d'ordinateur demande de nouvelles méthodes d'entreposage, de retrait et de préservation qui, à leur tour, représentent d'importants besoins de formation pour un personnel habitué à la gestion de documents plus tangibles.

Le groupe de travail se préoccupa également des questions touchant l'authenticité, l'intégrité et la sécurité des collections électroniques. Il doit y avoir une validation et une vérification automatiques du contenu des acquisitions des publications électroniques afin d'assurer leur ***authenticité***. La validation et la vérification automatiques d'une publication électronique incluent les mises à jour, l'adressage calculé ou le datage numérique de l'ensemble du document, le suivi de la périodicité et des mécanismes automatiques de comparaison des copies. Si possible, la BNC devrait acquérir et également stocker des publications électroniques en mode non compressé afin d'éviter toute perte de données causée par la compression. L'équipe de travail a reconnu toutefois qu'il y a des cas, surtout pour des objets audio et vidéo, où la compression et la décompression peuvent être effectuées d'une manière telle que la perte de données est indétectable par l'oeil ou l'oreille humaine. Ces cas devront être identifiés à partir de critères spécifiques.

La ***sécurité*** des publications électroniques comprend non seulement la protection du contenu contre les changements délibérés ou accidentels, mais comporte également des mécanismes de sauvegarde, de stockage de secours et des mesures en cas de sinistre. Les mécanismes actuels de sauvegarde et de reconstruction des bases de données et les procédures en cas de sinistre, actuellement en place pour les serveurs de la BNC, devraient être suffisants pour assurer la sécurité des collections électroniques.

On ne saurait sous-estimer l'importance des **normes** dans la gestion des publications électroniques. L'absence d'adoption généralisée de normes pour les publications électroniques rendla collecte, la fourniture d'accès et la conservation des publications électroniques beaucoup plus difficiles. Par exemple, l'intégration de renseignements d'identification normalisés dans les publications électroniques en faciliterait la recherche et l'utilisation au public cible. Les normes s'étendent, entre autres, aux signatures électroniques, au datage, au langage de balisage généralisé (Standard Generalized Markup Language) en vue de l'identification, du stockage et de la présentation des publications électroniques, au langage de balisage hypertexte (HyperText Markup Language) ainsi qu'à la définition du type de document (Document Type Definition). La BNC devra surveiller activement l'élaboration de normes pour les publications électroniques et devra faciliter et promouvoir l'adoption de normes appropriées en collaboration avec les éditeurs électroniques.

La question des «**versions**» ajoute un élément intéressant à la problématique d'une gestion appropriée d'une collection de documents électroniques. On pense qu'il y aura plus de versions d'une publication électronique qu'il y en a d'une publication imprimée. Si, pour une quelconque publication, la seule version disponible à la BNC est une version non normalisée, la BNC devra convertir la publication électronique en une présentation normalisée afin d'en assurer l'accès au plus grand nombre. De même, la BNC devra acquérir et conserver toutes les versions en « présentation neutre» (c'est-à-dire normalisée) des publications électroniques canadiennes afin d'en maximiser l'accessibilité.

Il n'existe pas en ce moment de support électronique reconnu comme support d'**archivage** de longue durée de sorte que la conservation des documents électroniques entraîne pour l'instant des coûts importants de transfert de l'information de ruban à ruban ou de disque compact à disque compact. Un système de stockage hiérarchique peut toutefois résoudre un certain nombre de problèmes. Ce système peut régénérer automatiquement les supports à intervalles réguliers et peut transférer automatiquement l'information d'un type de stockage à un autre (direct, quasi direct, hors ligne). Le système sauvegarde automatiquement les données qu'il contrôle. Le groupe de travail recommanda à la BNC de considérer un tel système pour l'entreposage des documents électroniques.

À plus long terme, la décentralisation des responsabilités relatives à la conservation des publications électroniques a été fortement recommandée par l'équipe de projet. La BNC entend étudier différentes options pour le stockage décentralisé des publications électroniques pour fins de conservation. La collaboration d'autres intervenants fédéraux semble offrir une piste d'intervention que la BNC pourrait privilégier.

Le support technique

Durant le projet pilote, en raison des limites financières, les infrastructures existantes et des «gratuiciels» ont été utilisés pour prendre en charge toutes les fonctions du *PPPE*, y compris l'acquisition, le catalogage, l'accès aux publications électroniques et leur présentation.

Il est évident que ce type de système sera insuffisant advenant un volume plus important de publications. La BNC se devra de recourir à un système de gestion plus perfectionné aussitôt que le volume de traitement l'exigera. Entre-temps, la BNC s'emploie à définir les spécifications

d'un tel système qui devra, entre autres, fournir des statistiques d'accès aux publications, automatiser les différents processus d'acquisition, supporter l'enregistrement et l'affranchissement des droits d'auteur,rationaliser le contrôle des publications en série, stocker et indexer les métadonnées, établir un lien dynamique entre la notice bibliographique et le titre enregistré dans le serveur, donner accès par le WWW, comporter un important logiciel de conversion ainsi qu'un logiciel de compression et de décompression.

Le rôle de la BNC dans la création et la diffusion de documents numérisés

Comme on l'a vu par tout ce qui précède, la gestion des publications électroniques présente un défi d'envergure aux bibliothèques en général et à la BNC en particulier. Les défis particuliers de la BNC relèvent de questions associées au dépôt légal et à son rôle national de préservation et d'accès au patrimoine canadien. Un autre défi de taille consiste dans le rôle que les institutions comme la BNC se doivent de jouer dans le développement du contenu canadien véhiculé sur l'inforoute.

Il y a carence de contenu canadien sur l'inforoute comme en a fait état le Conseil consultatif sur l'autoroute de l'information lors de ses délibérations, dans son rapport et dans ses recommandations. Les questions de souveraineté culturelle et de la menace que représente notre voisin du sud, ne serait-ce que par le volume même de ses productions, y ont été l'objet de maintes prises de position. L'intervention et la stimulation de la part du gouvernement et du secteur privé sont essentielles si on veut assurer la production et la disponibilité de documents numérisés canadiens.

Plusieurs initiatives sont présentement en cours. J'aimerais mentionner ici les efforts d'Industrie Canada dans l'établissement du Réseau scolaire canadien, plus familièrement connu sous le nom de ResCol. Plusieurs millions de dollars seront investis dans la numérisation de matériel pédagogique et de produits d'intérêt général pour nos jeunes étudiants. L'emploi y sera également stimulé puisque la plupart des projets retenus font appel à la participation d'équipes de jeunes qui y développent de nouvelles capacités ainsi qu'un esprit d'entrepreneurship. D'autres institutions culturelles travaillent à la numérisation de leurs propres collections, comme par exemple le Musée canadien des civilisations et la Galerie nationale du Canada. Un grand nombre de ministères fédéraux ont développé des pages d'accueil qui donnent accès à leurs produits, publications et services. Les gouvernements à tous les niveaux voient dans cette nouvelle technologie une façon moderne de s'approcher des citoyens et de faire connaître leurs programmes et services.

Les institutions de recherche au Canada se préoccupent également de la faiblesse du contenu canadien accessible sur l'inforoute. Dix bibliothèques de recherche, incluant la BNC et l'ICIST, ont participé à l'élaboration d'un projet de bibliothèque virtuelle au Canada qui comprend des projets de numérisation d'importantes collections. Des aspects de recherche et de communication font également partie de la soumission du Groupe des dix. Le financement de ce projet[1] n'est toujours pas assuré quoique le projet ait reçu l'aval du Conseil consultatif et ait fait l'objet d'une recommandation spécifique. Des projets individuels de numérisation sont en cours dans plusieurs universités canadiennes notamment à l'Université de Toronto ainsi qu'à la Bibliothèque nationale du Québec. Le Conseil national de recherche travaille activement à une version électronique de certaines de ses publications.

En plus de participer activement, de concert avec d'autres agences culturelles, aux initiatives de numérisation d'Industrie Canada et d'avoir collaboré étroitement avec le Groupe des dix, la BNC cherche à développer l'expertise et l'infrastructure nécessaires pour participer au développement du contenu canadien, dans la limite de ses moyens. À cette fin, un Comité de numérisation a été établi, formé de représentants de tous les secteurs intéressés au sujet: curateurs, bibliothécaires de référence, spécialistes du balisage et de l'encodage, spécialistes des publications et du marketing. L'objectif est double: recommander des projets de numérisation de type et d'envergure différents et élaborer des politiques et procédures visant à encadrer la gestion de projets de numérisation. Des critères de sélection et une grille d'évaluation ont été développés et un appel d'offres lancé en septembre 1995. Des projets sont actifs depuis janvier 1996.

Les projets proposés doivent répondre aux deux objectifs suivants: promouvoir les services et les collections de la BNC et viser un marché bien défini. Le choix de projets sera fait de façon à maintenir un équilibre entre des projets de type thématique qui cherchent à présenter les collections de la BNC autour d'un thème et dans un emballage propre à attirer l'attention du grand public, des projets qui rendent disponibles sous forme numérisée des collections spécifiques et d'intérêt particulier pour les études canadiennes et des projets qui facilitent l'accès aux collections grâce à la numérisation d'outils spécialisés de recherche et de localisation d'information.

Quelques problèmes associés à la numérisation

Les expériences menées à ce jour ainsi que les renseignements obtenus de nos partenaires ont mis en lumière quelques points importants.

Le financement et les ressources

Avec les coûts de la numérisation estimés entre deux et six dollars la page, il est impossible de penser qu'il y aura des progrès substantiels dans le développement du contenu canadien sans qu'il y ait un afflux sérieux de ressources en provenance des secteurs privé et public. Il y a également des coûts importants en personnel associés à tout développement de produit électronique, que ce soit pour la sélection ou la préparation du matériel, l'affranchissement des droits d'auteur, l'encodage ou le balisage des textes. Toutes ces fonctions exigent une formation.

Les modes d'accès

Le développement d'outils de navigation est essentiel au développement et

1. «The Can-Linked initiative/Le projet canadien.» *Libris* (June/Juin 1995).

à l'exploitation des ressources diffusées sur l'inforoute. De bons outils de navigation favoriseraient l'utilisation des produits canadiens. Les outils de navigation qui existent actuellement sont limités et de qualité inégale. Assurer le type d' accès nécessaire à l'exploitation du potentiel de recherche et d'information qu'est Internet représente un important défi. Les bibliothèques sont particulièrement qualifiées pour répondre à ce défi.

Les normes

Un effort concerté est essentiel pour résoudre la question des normes dans ce typed'environnement en évolution constante.

L'infrastructure

La combinaison des différents médias conduit à des changements d'infrastructure de façon à ce que soient encouragées la création des produits et l'exploitation des services multimédias. Il y a compétition entre les industries du câble et du téléphone pour prendre le contrôle du marché. Les résultats de cette compétition auront un impact important sur l'accès, la qualité et les coûts des services .

Le droit d'auteur

Pour des institutions telles que la BNC, la question du droit d'auteur a un impact sur les possibilités de conversion de matériel pour l'inforoute. Même si les détenteurs de droit sont prêts à donner leur accord, l'affranchissement des droits est une activité spécialisée qui demande beaucoup de temps et est susceptible de créer d'importants délais dans la numérisation de certaines oeuvres et collections.

La conservation

On a tendance à croire que les produits électroniques vont éliminer ou réduire substantiellement les besoins de préservation des originaux. Cependant, alors qu'il est vrai que l'accès au contenu est grandement favorisé par la diffusion des oeuvres sur des réseaux tel Internet, le processus de conversion ne réduit pas les besoins en conservation des originaux; il crée même à cet égard des défis additionnels.

Le défi majeur: la sélection

Étant donné l'importance des collections et les coûts prévisibles de la numérisation, la sélection de ce qui doit être numérisé prioritairement est un problème capital. D'où l'importance de projets pilotes pour nous aider à déterminer ce que le grand public et les chercheurs veulent trouver dans Internet. On sait déjà que les décisions de sélection devront être faites en tenant compte de différents facteurs souvent en conflit les uns avec les autres. La valeur respective du produit pour la recherche académique, le grand public et les objectifs du gouvernement sont autant d'éléments dont il faudra tenir compte. La rareté, la fragilité ou l'accès limité des originaux sont également des éléments importants dans le processsus de sélection.

J'admets que c'est très peu sur un sujet si vaste et si complexe. J'aimerais toutefois ajouter, en guise de conclusion, que l'objectif de la BNC en s'engageant dans ce type d'activités est de créer une collection de Canadiana sur support électronique, un riche trésor pour l'étude et l'appréciation du Canada et des Canadiens. Grâce à la numérisation de notre mémoire collective, les Canadiens auront accès à des oeuvres actuelles, passées et futures dans un nombre grandissant de disciplines et de médias. Un tel trésor demande des outils de navigation sophistiqués et un engagement explicite à la préservation à long terme de l'information numérisée. Une telle vision demande un partenariat vigoureux des secteurs privé, public et à but non lucratif car elle doit pouvoir compter, pour devenirréalité, sur des penseurs et des planificateurs dynamiques et créatifs.

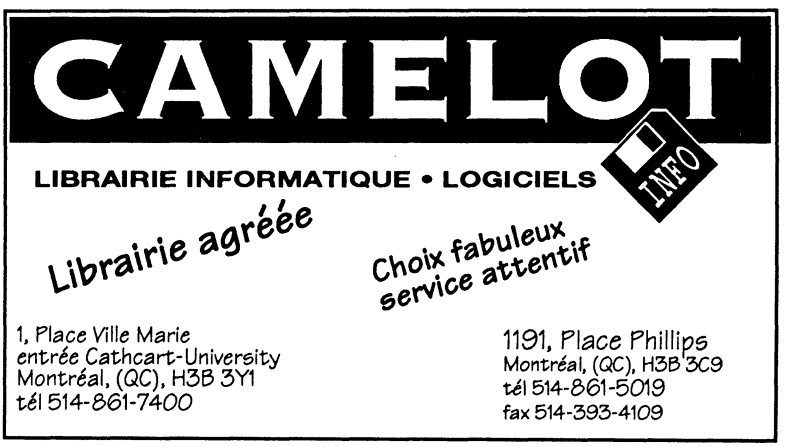
Sources consultées

Arms, William Y. K*ey architectural issues in the digital library*. Corporation for National Research Initiatives. Adresse électronique: http://www.cnri.reston.va.us/home/cstr/arch/slides.html

Bailey, Charles W., Jr. 1996. Network-based electronic publishing of scholarly work: a selective bibliography. *The Public-Access Computer Systems Review* 6 (1), Version 19, April 96. Adresse électronique: http://info.lib.uh.edu/pr/v6/n1/bail6n1.html

JEP: *The Journal of Electronic Publishing*. University of Michigan Press. Adresse électronique: http:// www.press.umich.edu:80/jep/

Jog, Vijay. June 1995. *Cost and revenue structure of academic journals: paper-based versus E-journals*. Submitted to Virtual Products, Industry Canada. 51p. Adresse électronique: http://schoolnet2.carleton.ca/english/biz/economics/vijayjog.html

Kahn, Robert. May 1995. Corporation for National Research Initiatives and Robert Wilensky, University of California at Berkeley. *A framework for distributed digital object services*. 15p. Adresse électronique: http://www.cnri.reston.va.us/home/cstr/arch/k-w.html

Workshop on issues in the field of national deposit collections of electronic publications: rapport sommaire. 1995 Luxembourg: European Commission's DGXIII/E. 9 p. Adresse électronique: http://www2.echo.lu/libraries/en/depo-rpt.html