

ÉTUDE

*Les documents électroniques et le calendrier de conservation ou les considérations d'un pauvre archiviste qui essaie de suivre l'évolution technologique**

Michel Lévesque

Qui possède un ordinateur?

Que faites-vous à chaque jour?

Moi, à chaque jour, j'ouvre mon ordinateur. Je consulte mes répertoires ou mes fichiers. Ou bien je consulte mon courrier électronique avec LOTUS. Ou bien je vais chercher sur Internet avec NETSCAPE. Je garde toujours en tête que je peux en un instant revenir à mes répertoires ou à mon courrier électronique.

À chaque jour, je crée un ou des fichiers en WORD PERFECT 6.1. Ou bien j'en lis certains et j'en corrige d'autres en enlevant ou en ajoutant des informations. Je sauvegarde en nommant mes fichiers selon une logique toute personnelle. J'imprime des dizaines de pages. J'en jette plusieurs mais j'en conserve aussi plusieurs.

Je peux aussi consulter ma banque de données COBA en gestion documentaire. Voir quelles données ont été enregistrées par la Direction régionale de Rimouski ou de Val D'Or. Je peux faire imprimer ma liste de mots-clés pour les valider.

* Ce texte est celui, remanié, de la conférence prononcée par l'auteur dans le cadre du XXVII^e congrès de l'Association des archivistes du Québec à Magog-Orford le 29 mai 1998

Et quoi encore...

L'informatique a ceci de merveilleux: je peux faire tout cela uniquement en appuyant sur quelques touches d'un clavier ou en jouant avec une souris.

Mais si je pose mon regard d'archiviste sur mes pratiques, qu'est-ce que j'en conclus?

Je crois gérer mon information efficacement, alors que dans les faits je ne gère pas grand chose. J'ai accumulé, souvent même détruit, des informations sans me préoccuper d'un plan de classification ou d'un calendrier de conservation. J'augmente mes capacités de stockage, et le tour est joué. Je me crée des fichiers d'archives et, comme le terme veut bien dire ce qu'il veut dire, je me crois en sécurité et bon gestionnaire.

Imaginez alors le même scénario lorsque plusieurs individus font la même chose dans leurs organismes.

Mon propos ne sera pas de démontrer l'étendue de la problématique qui provient de l'utilisation de l'informatique et de l'électronique par rapport à la discipline archivistique. Je vais plutôt essayer de centrer mon exposé sur les conséquences qu'elles ont sur le calendrier de conservation des documents. Que bouleversent-elles ou qui bouleversent-elles?

Aujourd'hui, les possibilités technologiques sont tellement grandes en matière de création, de reproduction et de traitement des documents dans cet univers qu'elles forcent l'archiviste à se poser ce genre de questions. Certains éléments de son calendrier de conservation qu'il croyait évidents dans le monde du papier doivent maintenant être abordés de nouveau. Les technologies de l'information avec leurs possibilités si grandes, posent à l'archiviste un lot de questions qui ne sont pas sans impacts sur la confection d'un calendrier de conservation. Ces questions seront abordées au fil de ma présentation. Elles tournent toutes autour des éléments principaux que l'on retrouve dans la confection d'un calendrier de conservation.

D'abord, fondamentalement, elles ont trait à la nature même d'un document dans un environnement électronique (base de données et bureautique), à sa valeur probante et à son cycle de vie, notamment au principe de semi-activité dans le monde électronique. Ensuite, les différents formats et les supports d'information sont des éléments complexes également à considérer. Une autre grande question qui sera abordée est celle de la sélection des documents dans les environnements électroniques. Comment peut-on en effet parvenir à réaliser celle-ci?

Ainsi je croiserai donc sur mon chemin ces différentes questions en traitant en premier lieu des documents numérisés, ensuite des banques de données et, pour terminer, des documents bureautiques.

LA NUMÉRISATION

La numérisation est un processus qui permet de convertir de l'information consignée sur support papier ou sur film – formulaires, rapports, cartes géographiques, correspondances, photographies, dossiers d'employés, etc. – sous un format électronique «image», c'est-à-dire la photographie électronique d'une information en code binaire.

L'objectif premier de la numérisation est de réduire les coûts d'entreposage des dossiers et des documents. Elle favorise aussi un repérage plus rapide de l'information. Pour certains, il s'agit seulement de se doter d'un support de consultation dans le but de protéger les documents originaux. Tandis que pour d'autres, il s'agit de substituer des «images» à des originaux, en conservant à ces images une valeur de preuve légale selon les obligations énoncées aux articles 2840 à 2843 du *Code civil du Québec*.

Ces «images» sont-elles de nouveaux documents? J'ai un nouveau support qui contient des informations déjà consignées sur un support original. Il s'agit d'une copie de l'information sur un autre support. Mais les deux supports doivent être considérés comme des exemplaires principaux parce qu'ils sont complémentaires ou qu'ils peuvent se substituer l'un à l'autre. Je serais donc tenté de conclure que j'ai un nouveau document. Comme celui-ci doit être évalué en fonction de l'information qu'il consigne et que celle-ci a déjà fait l'objet d'une évaluation, il suffit seulement de la valider.

Ceci explique pourquoi la règle de conservation spécifie les deux supports de conservation et montre ainsi le transfert de l'information. Je peux procéder à ce transfert dès le stade actif ou attendre au stade semi-actif. La règle de conservation devra refléter parfaitement mon choix. Dans le cas où le support numérique sert de support de consultation, il y a transfert d'information mais la règle de conservation sera identique à celle fixée pour les informations sur support papier.

En règle générale, l'établissement, et surtout l'application, des durées de conservation aux stades actif et semi-actif causent moins de problèmes pour des documents ou des dossiers gérés selon les années financières. Elles impliquent des durées fixes. Il sera donc plus facile de faire coïncider l'enregistrement sur disque des informations avec les durées de conservation retenues. À la limite, comme les durées actives ne comprennent pas normalement l'année en cours, il serait conseillé de procéder à la numérisation seulement lorsque l'année est terminée. Cette contrainte n'existe plus lorsque je procède à la numérisation au moment où les documents deviennent semi-actifs.

Il en va tout autrement lorsque je procède à la numérisation des documents et des dossiers gérés selon des durées de conservation ouvertes, tels les procédures, les dossiers de plainte, les dossiers de contrat, les avis juridiques, etc. Comment réaliser leur numérisation en gardant à l'esprit que des documents peuvent s'ajouter à chaque mois au dossier ou être remplacés par de nouvelles versions? Devrais-je attendre leur dénouement? Si oui, le problème est en grande partie résolu. Sinon, quels seront les mécanismes qui permettront de maintenir l'idée du dossier? Des liens entre les disques optiques? Des disques qui rendent possible l'ajout d'information au fur et à mesure qu'elle est créée, exemple: le WORM? Des index qui bénéficieront de la «puissance» des ordinateurs pour leur exploitation maximale au repérage et au couplage de l'information? Il faudrait m'éclairer. Mais le calendrier de conservation devrait pouvoir s'adapter aux moyens mis de l'avant pour contrer ces difficultés.

Pensez, entre autres, à la double difficulté que présentent les dossiers d'employés ou les dossiers d'étudiants. La règle de conservation est ouverte à l'actif et tributaire de l'année de naissance pour le calcul de la durée au semi-actif. Il faudra dans ces cas modifier les règles actuelles de conservation ou bien conserver longtemps

plusieurs documents répartis sur plusieurs disques. En ce sens, la numérisation permettra le requestionnement d'évaluations qui commencent peut-être elles-mêmes à prendre «de l'âge».

De plus, certains remettent en cause le principe de semi-activité. Est-ce justifié? Ce principe est basé sur la fréquence d'utilisation des documents et des dossiers. Numérisés ou non, lorsqu'ils sont utilisés occasionnellement, ils doivent être considérés comme semi-actifs. Par une évaluation des besoins administratifs, mais surtout légaux et financiers, cette durée de conservation semi-active est justifiée. La seule différence, et encore est-ce vraiment une différence, c'est que les disques optiques peuvent être conservés sur place, dans l'organisme, et non dans des entrepôts comme c'est le cas pour les boîtes de dossiers. Ils prennent quand même moins d'espace que ces boîtes.

Outre le but de fixer des durées actives et semi-actives aux documents et dossiers numérisés, il faut aussi déterminer un mode de disposition. La conservation et la destruction en totalité ne causent aucun problème. L'échantillonnage et la sélection compliquent la situation. Certains préféreront ne faire ni l'un ni l'autre et conserveront les documents et les dossiers visés. C'est moins contraignant mais peut-être plus coûteux à la longue.

Si je décide de procéder quand même à un échantillonnage, il faudra, dans la mesure du possible, que je prélève l'échantillon à l'avance et que je le place sur un disque autre que ceux sur lesquels se retrouvent les documents et dossiers à détruire. Cette méthode est plus simple. Techniquement, il est possible de prélever des informations déjà enregistrées sur un disque et de les transférer sur un autre disque. Mais quels coûts sont engendrés par cette opération? Quel traitement devrais-je faire pour accéder à nouveau aux informations?

Cette méthode serait idéale pour le mode de tri qu'est la sélection. Elle implique un choix subjectif de documents à l'intérieur d'un dossier. Puis-je le faire à l'avance? Quelle contrainte cela m'impose-t-il? Certains jugent qu'il vaudrait mieux éviter la sélection basée sur des choix subjectifs et conserver l'ensemble des dossiers. La numérisation leur donnera peut-être raison. Quoiqu'il en soit, un mode de disposition devra toujours être établi après une évaluation sérieuse de la méthode de tri retenue. De plus, il devra toujours spécifier l'information qui constituera l'échantillon, par exemples, les dossiers des individus nés aux années 1 et 6, le programme et le rapport de stage, 1 dossier sur 10, etc.

Dans la mesure où j'envisage la destruction des documents originaux, il va sans dire que l'établissement des règles de conservation à l'actif et au semi-actif vaudra la peine seulement si je respecte les dispositions prévues au *Code civil du Québec* concernant la reproduction des documents. Dans la même veine, je devrai utiliser un système d'imagerie ouvert et respecter les standards actuels pour pouvoir exploiter l'information numérisée. Une règle de conservation ne servirait à rien si l'information était perdue ou était illisible.

LES BANQUES DE DONNÉES

Ce sont véritablement de beaux cas pour les archivistes. Une banque de données est un ensemble intégré de données où celles-ci sont réparties dans un ensemble de fichiers et d'enregistrements logiquement reliés entre eux.

Les données, pour bien comprendre, représentent les concepts, c'est-à-dire l'information. Elles sont organisées en caractères (une lettre, un nombre, un symbole), puis en champs (plusieurs caractères logiquement reliés entre eux: NAS, code d'étudiant), puis en enregistrements (ensemble de champs reliés entre eux: toutes les informations relatives à un permis de conduire), puis en fichiers (ensemble d'enregistrements reliés entre eux: tous les permis de conduire émis). On utilise le code binaire pour représenter les données et les instructions de traitement: 0 et 1 (exemples de codes binaires: EBCDIC et ASCII).

À titre d'exemple, le terme «archives» se traduit en ASCII par la chaîne suivante:

```
10000011010010100001110010001001001101011010001011010011
```

Un système de gestion de bases de données structure habituellement les données à sa manière: il comprend un dictionnaire de données, un langage de requête et les données. On retrouve donc des banques de données hiérarchiques, en réseau et, les plus utilisées actuellement, relationnelles. C'est l'organisation la plus facile car il n'y a pas de chemin d'accès à travers une hiérarchisation de nœuds. Les éléments de données sont plutôt conservés dans diverses tables faites de rangées et de colonnes. Les données sont repérées grâce à des index. Une table, qu'on appelle relation, ressemble à un fichier d'information sur un même sujet.

Pour accéder à ses données, il faut un logiciel et un ordinateur (appareil). C'est pourquoi plusieurs spécialistes parlent de l'opacité de l'information. Elle ne peut être lue que par une machine. Voilà donc un gros problème. Et ce n'est pas le seul.

La question que les spécialistes se posent actuellement est: sommes-nous en présence d'un document? En fait, je retrouve dans une banque de données de l'information brute qui, lorsque traitée, fournit des documents, tels des listes, des rapports, des formulaires, des statistiques, des tableaux, etc. Or, sans vouloir produire un document en tant que tel, je peux uniquement consulter la banque pour connaître telle ou telle information, au même titre que je me servais d'un index sur support papier ou d'un registre. C'est d'ailleurs le rapprochement que le Comité sur les documents électroniques du Conseil international des archives faisait dans son étude publiée en février 1997.

Par le langage de requête et des panoramas préétablis, je peux effectivement accéder à l'information consignée dans la banque. J'aurais donc tendance à considérer la banque de données comme un document qui peut emmagasiner un ensemble impressionnant d'informations. Encore ici, si je sais lire, l'étude des experts du CIA parle d'un *compound record*, bien qu'ils fassent pour en arriver là de multiples nuances et compromis.

Par contre, certains ne considèrent pas la banque de données comme un document. Ils invoquent l'absence de structure, telle qu'elle se retrouve dans un document

textuel sur support papier ou informatique. Or, n'en déplaise à plusieurs, la structure logique d'une banque de données constitue la démonstration sans équivoque que les données sont organisées de manière à pouvoir les utiliser ou les consulter sans contrainte, si ce n'est celle de l'utilisation d'un logiciel et d'un ordinateur. La structure de l'information sur bande vidéo ou sonore a-t-elle fait autant jaser dans le passé?

En corollaire, les opposants prétendent qu'une banque de données n'a pas de valeur probante. Cela m'apparaît curieux. Je crois au contraire qu'une banque de données atteste des activités d'un organisme et qu'elle peut donc avoir aussi une valeur probante. Je ne crois pas qu'il s'agisse d'une *data library*. Les données consignées sont relatives à des fonctions et des activités menées par l'organisme. Pensez entre autres aux banques de données de la RAMQ, du ministère du Revenu ou d'Hydro-Québec.

Nonobstant cette problématique, les détenteurs principaux des banques de données dans un organisme alimentent ces banques. Ils y ajoutent des informations. Ils en détruisent par écrasement de données. Ils en déplacent sur des supports plus économiques ou lors de refontes de systèmes informatiques. Plusieurs archivistes trouvent que cette raison est suffisante en elle-même pour que les banques de données soient gérées de la même manière que le sont les documents et les dossiers sur support traditionnel.

Ainsi, il faut déterminer pour chaque banque de données, pour les fichiers ou les enregistrements le cas échéant, des durées de conservation et des modes de disposition. La règle de conservation devra comprendre d'autres éléments descriptifs. La description devrait aussi être reliée au système informatique dont la banque fait partie. Des exemples de formulaires existent, dont entre autres ceux des Archives nationales du Québec et de la Ville de Montréal.

En règle générale, les données actives seront toujours conservées sur disque rigide. Tandis que les données semi-actives seront transférées sur cartouches magnétiques ou sur disques optiques. Par contre, depuis quelques temps, je vois des organismes conserver sur disque rigide, dans des fichiers séparés, les données semi-actives. Les règles de conservation devront tenir compte des supports selon leur stade de vie.

La notion de semi-activité préoccupe encore plusieurs spécialistes. Mais que se passe-t-il réellement? Il devient presque évident que les données perdent de leur acuité d'année en année. Le principe même de la constitution des banques de données serait remis en cause s'il en était autrement. L'information qui doit être active est bien celle actuelle et constamment renouvelée, à moins que l'organisme consacre ses énergies à des études rétrospectives. Et, encore là, cela ne nierait en rien le concept de semi-activité. Le lieu d'entreposage, les informaticiens parleraient d'archivage, peut être autre qu'un magasin ou un entrepôt. La nature des données restera la même. Les informaticiens viennent d'introduire un nouveau concept de gestion: l'entrepôt de données, le *data warehouse*. Nos chemins convergeraient-ils?

Plusieurs expériences dans l'élaboration des règles de conservation conduisent à des typologies de fichiers. Je peux donc élaborer des règles de conservation générales pour un ensemble de fichiers, tandis que j'en fixerai des plus spécifiques selon les données enregistrées. À titre d'exemples, je prends certaines règles élaborées par deux ministères. Ainsi, je retrouve dans leurs calendriers des règles générales pour les fichiers suivants:

FICHIERS	RÈGLES GÉNÉRALES
Fichiers de développement	888*- 1 - D * Conserver selon les besoins du développeur
Fichiers de bibliothèques	999 - 1* - D * une génération
Fichiers transitoires/temporaires Fichiers de travail Fichiers de test Fichiers d'entretien Fichiers de liaison	888* - 0 - D * Durée de vie du traitement
Fichiers d'entrée	888* - 0 - D * Tant qu'utiles
Fichiers statistiques du système d'exploitation	888* - 0 - D * Tant qu'utiles
Fichiers personnels	888* - 0 - D * Tant qu'utiles

Il faut aussi tenir compte que les banques de données peuvent être cumulatives, non cumulatives ou mixtes. Ces catégories influenceront les durées de conservation actives et semi-actives que je fixerai.

Le principe à la base de tout calendrier de conservation demeure l'évaluation. Avec l'aide des créateurs et des utilisateurs des données, il faut évaluer combien de temps elles leur sont utiles. Par la suite, il faut faire une évaluation archivistique de l'intérêt qu'elles peuvent représenter pour la recherche historique et scientifique. Je peux donc rester à un niveau plus général en prenant la banque de données dans son ensemble. Je peux aussi être plus spécifique et considérer des parties de celle-ci, soit certains fichiers et même certains enregistrements.

Ainsi, je peux déterminer que, dans la banque de données non-cumulative X au stade actif, certains fichiers doivent être conservés 5 ans et que certains enregistrements le seront jusqu'à remplacement. Je pourrai aussi avoir une banque de données cumulatives où toutes les données seront conservées à l'actif 2 ans et transférées au semi-actif pour une durée de 5 ans. La règle de conservation devra toujours traduire les résultats de l'évaluation.

Selon ce que j'ai vu jusqu'à maintenant, les organismes ont surtout inscrit des durées de conservation pour les fichiers maîtres et des fichiers historiques. Bien que les banques de données décrites soient mixtes, les règles ne rendent pas toujours compte de cette réalité. Peu de codes 999 ou d'expression «jusqu'à remplacement par de nouvelles données» ont été employés. Certains abusent même de la durée ouverte avec comme remarque «durée de vie du système». Cette durée devrait être une exception. Je serais tenté de croire que parfois les organismes ne vont pas suffisamment loin dans leur évaluation. Tout comme l'évaluation archivistique s'avère parfois difficile et conclue à la conservation de banques de données entières. Cela est un moindre mal, mais impliquera une réévaluation lorsque les banques seront versées.

Mais, avant tout, puis-je conserver en permanence des données? Le problème majeur, comme je l'ai mentionné, est la dépendance de la banque de données par rapport au logiciel et à l'appareil qui permettent de l'exploiter. L'autre problème qui préoccupe beaucoup les archivistes est la longévité du support.

Une étude récente des Archives nationales du Québec démontre la faisabilité de:

1. la conversion technique des documents informatiques existants sur les plates-formes maxi, mini et micro-ordinateurs vers une technologie unique qui soit la plus indépendante possible de l'évolution technologique;
2. la conversion automatisée ou semi-automatisée à l'aide d'outils déjà existants (IDMS, ORACLE) et la conversion des autres technologies par des programmations spéciales;
3. la conversion des différents types de technologies de banques de données (réseau, relationnelle, pseudo-relationnelle, hiérarchique) vers un format séquentiel unique;
4. du maintien, suite à la conversion technique, de la cohérence et de l'intégrité des données même lors d'extractions/conversions partielles;
5. la conversion de la définition physique des données pour certaines technologies;
6. la possibilité de lire et d'exploiter les données converties.

Ces conclusions devraient minimalement rassurer ceux qui s'inquiètent de la conservation à long terme des données. Certaines questions, telles celles des coûts périodiques de l'entretien des données, des reconversions futures, de la conversion des dictionnaires de données, de la micro-évaluation des données, etc., restent en suspens.

La question du support demeure aussi pertinente. La cartouche magnétique est le support privilégié. Mais, dans certains milieux, l'intérêt va au disque optique. Or, quelques institutions sont plus prudentes et optent dans l'immédiat pour l'utilisation du support COM, donc un format analogique.

Il faut, je pense, que j'apporte les nuances qui s'imposent. Il peut être effectivement conseillé d'utiliser le COM pour des fichiers de données agrégées qui constituent un «portrait» résultant d'une application donnée, exemples les états financiers, les listes statistiques, etc. Il ne servirait à rien de conserver une banque de données comptable pour traiter les données à d'autres fins que celles initialement prévues. Et comme il peut en coûter plusieurs milliers de dollars en développement de programmation, il

est peut-être mieux de conserver des fichiers traités ayant une valeur permanente. Ainsi, il pourrait y avoir seulement un «transfert» vers le support mentionné. Les chercheurs sont-ils intéressés à payer pour faire de nouvelles programmations? D'ailleurs, il faudrait aussi vérifier si ces fichiers n'ont pas donné lieu à des documents sur support papier.

À cause de sa longévité, le microfilm est un support sûr. Pas besoin, non plus, de repiquer l'information à intervalle fixe; même que l'information pourrait être retransformée en format numérique, donc informatiquement réexploitable. Or, à cause de l'arrivée de formats compatibles avec plusieurs environnements informatiques, les données pourraient facilement être transférées sur disque optique. L'avantage est qu'elles resteraient numériques.

Dans le cas de banques de données où j'évalue que les données pourraient servir aussi à la recherche, il devient plus logique de les laisser sur support numérique ou magnétique, converties en fichiers séquentiels selon un format standard. Elles permettraient de donner une image de ce que l'organisme faisait mais elles pourraient aussi servir à d'autres fins qu'au témoignage.

Ces approches ont bien sûr des conséquences sur les règles de conservation relatives aux données conservées en permanence, et même pour celles reliées à de longues périodes semi-actives. Si l'organisme choisit le COM, il devra l'indiquer pour les fichiers spécifiques sur lesquels cela s'appliquera, ou bien il pourra faire une règle générale où il spécifiera que les fichiers à conservation permanente seront transférés sur ce support. Les modes de disposition, soit la conservation et la destruction, seront associés respectivement au COM et au disque rigide ou au ruban magnétique. De la même façon, l'organisme, qui procédera au transfert des données du disque rigide au disque optique ou au ruban magnétique, devra indiquer clairement les deux supports. Ici encore, la conservation et la destruction devront être jumelées respectivement au ruban magnétique ou au disque optique et au disque rigide.

Tout comme pour la numérisation, la destruction et la conservation en totalité des données d'une banque ne causent pas de problème. Mais l'échantillonnage ou la sélection sont-ils possibles et conseillés? L'expérience des ANQ montre que des extractions de données dans les banques sont réalisables. Elle mentionne qu'il est plus facile par contre de convertir une banque en entier plutôt que des parties de celle-ci. Je comprends que le risque de perdre des données et surtout des liens renforce cet argument. Or, la règle de conservation devra correspondre à la décision d'utiliser ou non ces types de tri.

Avec une approche proactive, la situation serait différente. En intervenant dès la création des banques de données et en suivant leur évolution, il serait possible de déterminer dans quelle banque se retrouvent les données historiques. Je pourrais donc demander la création d'un fichier qui contiendrait des données converties. Ce fichier pourrait s'appeler «fichier d'archives». Il serait transféré au dépôt d'archives selon une périodicité variable d'une banque à l'autre et mis à jour sur une base continue. Une règle de conservation générale serait alors suffisante pour gérer ces fichiers.

Les banques de données ne sont compréhensibles qu'à l'aide de la documentation du système et du dictionnaire de données. Il faut donc obligatoirement faire une

règle de conservation générale pour ces outils en tenant compte de celles établies pour les banques de données. Ces outils devront être conservés tant et aussi longtemps qu'ils réfèrent à des banques. Ceci est particulièrement vrai lorsque certaines banques, ou parties de celles-ci, seront versées. Généralement, il est ajouté au mode de disposition de ces outils une remarque spécifiant que la documentation de système et le dictionnaire des données doivent être versés s'ils documentent une banque de données elle-même versée. Malheureusement, ces documents sont parfois inexistantes ou incomplets. Si cela affecte la banque, elle ne devra pas être conservée.

Les métadonnées, c'est-à-dire les données sur les données, sont peut-être la solution. Chaque banque ou chaque fichier devrait être accompagné de ses métadonnées. Ainsi, ces dernières pourraient être conservées avec les banques et fichiers à conservation permanente. Les métadonnées devraient être gérées par une règle de conservation similaire à celle des dictionnaires de données et de la documentation de système, sauf si l'ensemble est sur support numérique. Une seule règle serait alors suffisante.

LES DOCUMENTS BUREAUTIQUES

Je ne savais vraiment pas si j'allais aborder le cas des documents «bureautiques» ou des «documents textuels sur support informatique». Comme la gestion des documents sur support papier avait été dans le passé un défi pour l'archiviste, la gestion des documents créés à l'aide des logiciels de traitement de texte ou des chiffriers électroniques en représente un nouveau de taille.

Pensez à mon scénario du début. Avec l'arrivée de la micro-informatique, tous ont accès à l'ordinateur dans leur travail quotidien. Chaque personne crée, dans son environnement informatique, des lettres, des rapports, des études, des tableaux, des notes, etc., qu'elle sauvegarde dans des fichiers et des répertoires, souvent comme elle le veut, sans la plupart du temps les associer à des cotes de classification ou à des durées de conservation. Si bien que des milliers et des milliers de fichiers sont conservés selon le bon vouloir de chacun. Certaines règles d'archivage, pour reprendre l'expression des informaticiens, existent, mais elles ne sont pas véritablement des règles de conservation qui permettent une gestion efficace et réfléchie des disques et des disquettes, basée sur l'évaluation de la valeur de l'information.

Il faut comprendre que la situation n'est pas très évidente. Est-ce que toute cette production de fichiers n'est qu'une production de fichiers de travail qui ne contiendraient que des documents secondaires; ces derniers ayant été imprimés et classés dans les différents postes de classement? Si tel était le cas, mon rôle d'archiviste serait de gérer ces fichiers et ces répertoires en appliquant des règles de conservation conçues pour les exemplaires secondaires.

Or, une tendance contraire émerge dans certains organismes. Ils considèrent qu'il vaut mieux gérer les documents «bureautiques» que des dossiers et des documents sur support papier en classeurs ou dans des boîtes. Pourquoi conserver effectivement en double l'information?

Un premier obstacle auquel sont confrontées les organisations est le caractère légal du document «bureautique». Plusieurs craignent de conserver des documents qui ne seront pas reconnus légalement. Or, dans des travaux concernant l'échange de documents informatisés et le commerce électronique, les auteurs Trudel, Lefebvre et Parisien, de la Faculté de droit de l'Université de Montréal, concluent que:

Le *Code civil du Québec* [...] contient des dispositions spécifiques relativement aux inscriptions informatisées. Ces descriptions ne sont toutefois applicables qu'aux actes juridiques. On se rappellera qu'on entend par acte juridique «une manifestation de la volonté individuelle dans le but direct et immédiat de produire des effets juridiques, soit de créer, modifier, étendre, transférer, confirmer ou reconnaître un droit». Cette définition englobe le contrat, la renonciation à un droit, le paiement, la promesse d'achat ou de vente, etc. À cet égard, les dispositions au *Code civil du Québec* accordent une pleine reconnaissance en preuve aux documents qui reproduisent les données d'un acte juridique inscrit sur support informatique. Il s'agira de démontrer que le document est intelligible et qu'il «présente des garanties suffisamment sérieuses pour qu'on puisse s'y fier».

Les dispositions du *Code civil du Québec* posent toutefois un problème d'interprétation. D'aucuns prétendent qu'elles ne s'appliquent qu'aux transactions conclues directement sur support informatique. En conséquence, les parties devraient se méfier des transactions lorsque la volonté des parties s'est d'abord manifestée oralement ou par écrit et que l'inscription sur support informatique n'est intervenue que postérieurement. Il est incertain qu'une telle transaction soit visée par les dispositions du *Code civil du Québec* sur les inscriptions informatisées. Afin de résoudre ce problème, les parties devraient donc adopter une convention relative à la preuve.

Lorsqu'on est en présence d'un fait juridique plutôt que devant un acte juridique, il est clair que les dispositions du *Code civil du Québec* relatives aux inscriptions informatisées ne s'appliquent pas. On se souviendra qu'un fait juridique est un événement autre qu'une manifestation de volonté. Il s'agira par exemple d'un relevé. Relativement aux faits juridiques, la situation n'est toutefois pas alarmante. En effet, les dispositions générales du *Code civil du Québec* sont suffisamment souples pour permettre d'apporter en preuve les documents reproduisant les données inscrites sur un support informatisé. (Trudel, Lefebvre et Parisien 1993)

Il s'agit évidemment d'interprétations qui sont partagés par plusieurs. Beaucoup demeurent prudents et attendent la jurisprudence. Or, il est de plus en plus reconnu qu'en assurant par des moyens techniques l'intégrité et la fiabilité de l'information, les documents bureautiques conservent, s'il y a lieu, leur valeur légale ou, plus largement, leur valeur de preuve.

Je dois aussi aborder la notion du «dossier». Certains documents, bien qu'ayant en soi un sens, font doublement sens, un peu comme une valeur ajoutée, reliés à d'autres documents. C'est pourquoi ils sont regroupés ensemble dans une même chemise. Elle constitue un dossier, au sens de la définition du *Petit Robert*, c'est-à-dire «un ensemble de pièces relatives à une affaire...». Pour certains, tel Bearman, cette notion est fondamentale.

Les transactions (actions menées «à travers un objet») sont par définitions des actions communiquées d'une personne à une autre, d'une personne à une réserve d'information, comme un meuble à dossiers ou une base de données sur ordinateur, ou enfin d'une réserve d'information à une personne ou à un ordinateur. Ainsi, les transactions doivent émaner de l'esprit humain ou de la mémoire de l'ordinateur et être dites, écrites ou lues par le récepteur. Les dossiers sont les porteurs, les produits et la trace de la transaction. (Bearman 1996)

Je crois, en un certain sens, que cette notion est incontournable et qu'elle doit être prise en compte. Mais je crois qu'il faudrait ne pas strictement limiter l'attribution de la valeur probatoire aux dossiers, et, qui plus est, seulement si ces dossiers sont accompagnés de métadonnées.

Un deuxième obstacle est l'absence de classification des documents bureautiques. Les spécialistes s'entendent tous pour dire que les documents «bureautiques» doivent être classés selon un plan de classification comparable, voir le même, à celui servant pour les documents sur support traditionnel. Ils précisent que les documents bureautiques ne sont pas désincarnés parce qu'ils sont sur un autre support et qu'ils sont eux aussi le reflet des fonctions et des activités de l'organisme. Mme Carole Saulnier, dans un intéressant article qui vient d'être publié dans la revue *Archives*, va plus loin et démontre le bien-fondé pour une organisation d'avoir un «espace» commun où seraient obligatoirement classés, selon un plan de classification l'ensemble des documents administratifs, en opposition avec un «espace» personnel que chaque individu possède pour créer et conserver temporairement les documents avant qu'ils ne deviennent officiellement administratifs. De plus, elle propose une normalisation des noms de fichier, basée sur deux caractères du type de document produit, par exemple: AC pour Avis de convocation, MM pour Mémos, PV pour Procès-verbaux, etc. (Saulnier 1997-1998) Cette normalisation a le mérite de donner les prémices d'une solution qui, à mon sens, est partielle parce qu'elle peut limiter à la fois la classification à l'intérieur des sous-répertoires et le repérage de l'information. Tout en gardant cette normalisation, j'essaierais d'y adjoindre d'autres éléments classificatoires reliés aux sujets des documents.

Le troisième obstacle à la conservation des documents bureautiques est la multitude de formats informatiques existants. M. Yves Marcoux, dans un article paru en 1994 dans la revue *Archives*, citait que «l'information accessible par l'internet se retrouve dans plus de 300 formats différents.» (Marcoux 1994, 91) Il est d'ailleurs l'un des premiers à avoir parlé de l'opacité de l'information électronique en la définissant par la «nécessité d'une concordance parfaite entre le format d'un document électronique et celui auquel s'attend un système de traitement des documents» (Marcoux 1994, 91). Il précise en plus que

L'opacité de l'information électronique est très analogue au fait qu'il doit y avoir concordance parfaite entre un support de stockage et l'équipement de lecture utilisé. [...] cependant [...] il s'agit de deux phénomènes distincts: même si on possède un lecteur permettant de lire parfaitement une certaine disquette informatique, les documents électroniques enregistrés sur cette disquette peuvent malgré tout demeurer illisibles, si on ne possède pas de logiciel capable d'interpréter correctement le format dans lequel ils sont représentés.

... La plupart de formats de traitement de texte actuels sont des formats «propriétaires», c'est-à-dire sous le contrôle exclusif de producteurs de logiciels. Les producteurs sont donc libres de publier ou non les spécificités du format utilisé par leur logiciel, et de faire évoluer le format comme bon leur semble, d'une version de logiciel à une autre. (Marcoux 1994, 92)

À cause de ces raisons, la conversion d'un format à un autre est difficile, voire problématique et il ne s'agit encore que de documents textuels. Imaginez alors ce que réserve le monde des documents visuels, sonores ou multimédia.

La prolifération de ces différents formats sur Internet, l'échange de documents électroniques et le commerce électronique ont amené plusieurs personnes à s'interroger. Il est donc apparu des formats normalisés. Le SGML et le HTML en sont des exemples. Mais, il est de plus en plus question du XML, c'est-à-dire un langage extensible de balisage de documents. Il s'agit d'un SGML simplifié. Ce format est probablement, selon les spécialistes, celui qui s'imposera pour les années futures. Il permet de décrire la structure logique du document, la sémantique du contenu, à l'aide de balises. C'est un langage neutre et ouvert, indépendant des fournisseurs, des logiciels et des plateformes. Il permet l'interopérabilité entre les applications. Il utilise le jeu de caractères Unicode (ISO 10646). Ce format pourrait donc représenter un élément de solution à la conservation et à l'exploitation des documents bureautiques.

Volontairement, parce que je considère qu'il s'agit des plus importants, je me limiterai à ces trois obstacles pour lesquels des éléments de solutions légaux, archivistiques et informatiques concourent à assurer une pleine reconnaissance aux documents bureautiques.

Ainsi, dois-je faire de règles de conservation? Avec qui vais-je travailler?

Pour réaliser un calendrier de conservation des documents sur support papier, je dois premièrement bien connaître les documents créés par mon organisme. Je procède ensuite avec les utilisateurs à l'évaluation des valeurs administratives, légales, financières et historiques de ces documents. Je traduirais les résultats de cette évaluation par des règles de conservation pour chaque série de dossiers ou, le cas échéant, de documents. Suis-je en mesure de faire la même chose pour les documents bureautiques? Il me semble que oui. Dans la mesure où les dossiers sont «virtuellement» identifiables, il m'est possible de procéder à leur évaluation, et, le cas échéant, à l'évaluation des documents qu'ils contiennent.

Par rapport à l'évaluation des documents sur support papier, l'évaluation des documents bureautiques pourrait être différente. Je pourrais choisir d'évaluer document par document, même si, pour la valeur probante, certains documents doivent être constitutifs d'un seul dossier. Par exemple, je pourrais évaluer chaque document d'un dossier d'agrément d'un service d'archives privées et fixer 5 règles de conservation. Il est vrai que cette approche est contraignante mais elle répond aussi aux vœux de certains archivistes de mieux gérer des dossiers qui deviennent vite très épais. Un compromis acceptable est de procéder à ce genre d'évaluation mais de s'en tenir à la constitution, lorsque cela est vraiment nécessaire et produit un effet certain d'un meilleur contrôle du dossier, d'ensemble de documents appartenant à un même dossier et pour lequel des règles seraient élaborées. Je pense entre autres aux dossiers des employés. Une vision plus traditionnelle peut aussi être employée. Faire une règle de conservation pour l'ensemble des documents constitutifs du dossier. Je devrais ainsi être en mesure d'associer les règles de conservation aux répertoires, sous-répertoires ou même aux fichiers. Je pourrais faire coïncider alors les règles de conservation des documents sur support papier à celles des documents bureautiques.

Je devrai, par contre, garder l'esprit qu'il me faudra faire des liens entre les dossiers physiques et les dossiers «virtuels». Malgré l'arrivée du courrier électronique et du *workflow*, il y aura encore pour un certain temps cohabitation des deux types de dossiers. Ils sont complémentaires.

La réussite de la gestion des documents bureautiques dépendra des automatismes informatiques que je devrai obtenir afin de faire faire par la machine et le logiciel le repérage des délais associés aux fichiers, aux sous-répertoires et aux répertoires et leur application de façon intelligente. Même l'échantillonnage et la sélection seraient facilement exécutables par ce moyen. Des clefs codées associées aux fichiers permettraient de procéder au calcul des durées et au transfert des documents aux trois stades de vie.

L'évaluation des documents doit se faire obligatoirement avec les créateurs des documents, et parfois même avec les utilisateurs. Ils sont en mesure de dire à quoi servent les documents et combien de temps ils en ont besoin. Il est un peu paradoxal de constater que les systèmes informatiques sont faits pour eux mais ils ne sont pas souvent consultés quant aux durées de conservation des données et des documents. L'informaticien sera sans doute la personne la plus sollicitée. Il doit comprendre que les données et les documents doivent être évalués et il doit être en mesure de développer ces mécanismes qui rendront possible l'application des règles de conservation.

La notion de détenteur principal et de détenteur secondaire sera la plus difficile à appliquer. Je ne crois pas qu'elle soit foncièrement différente lorsqu'il s'agit de la gestion des documents bureautiques. Les rôles et mandats de chaque unité sont habituellement déterminés. C'est plutôt lorsque l'information se retrouvera dans un «espace» commun que les risques de confusion peuvent survenir. Il faudra alors prévoir associer de façon précise les fichiers, les sous-répertoires et les répertoires aux unités créatrices. Elles seules doivent pouvoir modifier les documents, tandis que les autres unités pourraient n'avoir qu'un droit de consultation. Cela impliquera par contre des clefs codées pour le respect de la confidentialité de certains documents.

EN CONCLUSION

Et l'intranet?

Et l'internet?

Et ça recommence...

Rassurez-vous, je ne recommencerai pas. Comme vous venez de le lire, je n'ai pas trouvé toutes les solutions. Il faut encore approfondir. Mais une chose est certaine: je crois que le calendrier de conservation des documents n'est pas prêt de devenir un document inactif. Au contraire. Il faudra parfois faire des adaptations mais il demeure un outil essentiel. Dans tous les sens du terme, il est bien cet élément de soutien de tout programme de gestion des documents.

Un calendrier de conservation demeure un calendrier. Je l'ai déjà dit, même dans le monde électronique son principe de base demeure l'évaluation. Toutefois, une adaptation du calendrier est nécessaire face aux moyens et techniques informatiques mises de l'avant. Ceci nous pousse à un questionnement de notions telles que la semi-activité.

Plusieurs expériences sont entreprises actuellement quant à l'élaboration de règles et moyens de conservation. Il y a donc de l'espoir. Je crois qu'une approche proactive de ces problèmes nous mènera plus loin sur la piste des automatismes néces-

saires à une application satisfaisante des calendriers de conservation dans un environnement électronique.

Deux facteurs sont essentiels pour parvenir à des solutions concernant l'élaboration et l'automatisation des règles de conservation. D'un côté le coût de ces solutions en regard des exigences de conservation et de l'autre la concertation essentielle entre l'archiviste et l'informaticien. Ceci est un facteur de succès incontournable pour la réalisation de règles de conservation pour les documents informatiques. L'archiviste et l'informaticien travaillent pour des individus qui font littéralement «vivre» leur organisation. Il faut donc coordonner les actions de chacun vers un même but: permettre de mieux vivre et de laisser des traces significatives de ce mieux vivre.

Michel Lévesque Chef du Service de la gestion documentaire au Secrétariat du Conseil du Trésor du Gouvernement du Québec.

BIBLIOGRAPHIE

- ARCHIVES NATIONALES DU QUÉBEC (ANQ). 1998. *Étude de faisabilité de la conversion technique des documents informatiques*. Rapport final, 18 janvier. Document non publié.
- ARCHIVES NATIONALES DU QUÉBEC (ANQ). 1991. *Modèle pour l'inscription des documents informatiques au calendrier de conservation: Guide explicatif*. Québec: Ministère de la Culture et des Communications.
- BEARMAN, David. 1996. Les archives virtuelles. In *XIII^e Congrès international des archives* (deuxième session plénière, troisième rapport subsidiaire). Beijing.
- INTERNATIONAL COUNCIL ON ARCHIVES – CONSEIL INTERNATIONAL DES ARCHIVES. COMMITTEE ON ELECTRONIC RECORDS. 1997. *Guide for Managing Electronic Records from an Archival Perspective*. Études CIA 8. France, Le Conseil, février.
- COUTURE, Carol. 1996. Le concept de document d'archives à l'aube du troisième millénaire. *Archives* 27, no 4: 3-19.
- DEMERS, Jean-Maurice. 1996-1997. La reproduction de documents pour fin de preuve selon le Code civil du Québec. *Archives* 28, nos 3-4: 3-21
- DLM-FORUM. 1997. *Guideline on best practices for using electronic information. How to deal with machine-readable data and electronic documents*. Luxembourg: Office for Official Publications of the European Communities.
- GRIMARD, Jacques. 1996. Gérer la préservation à long terme des archives électroniques ou préserver le médium et le message. *Archives* 27, no 4: 21-34.
- LEBEL, Marc. 1995. Méthodologie pour le calendrier des délais de conservation des documents informatiques. *Archives* 27, no 2: 51-73.

- MARCOUX, Yves. 1994. Les formats de documents électroniques en archivistique. *Archives* 26, nos 1-2: 85-100.
- MARCOUX, Yves. 1997. Possibilités de l'information électronique pour l'archivistique. In *Actes du XXVI congrès de l'Association des archivistes du Québec*, Aylmer, 28-31 mai: 117-122.
- PÉRIAT, Nicole. 1997-1998. Politique de gestion du courrier électronique: des mesures à prendre. *Archives* 29, no 1: 3-56.
- PHILLIPS, John T. 1997. What's in that Data Warehouse ? *Records Management Quarterly* 31, no 2: 54-56.
- PROVINCE DU QUÉBEC. ASSEMBLÉE NATIONALE DU QUÉBEC. 1991. *Code civil du Québec*. Québec: Éditeur officiel du Québec.
- SAULNIER, Carole. 1997-1998. Prolégomènes à la gestion des documents administratifs électroniques. *Archives* 29, no 1: 57-73.
- THIBAUDEAU, Kenneth. 1996. La recherche internationale: évaluation et conservation des bases de données scientifiques. In *XIII^e Congrès international des archives* (quatrième session plénière, troisième rapport subsidiaire). Beijing.
- TRUDEL, Pierre, Guy LEFEBVRE et Serge PARISIEN. *La preuve et la signature dans l'échange de documents informatisés au Québec*. 1993. Québec: Publications du Québec.
- VIGNEAU, André. 1996. Les documents informatiques: pour une classification efficace. *Archives* 27, no 3: 29-51.