

## Informations et réflexions

### Aperçu sur les problèmes d'archivage de données électroniques dans le cadre communal

par Frédéric Sardet, archiviste de la Ville de Lausanne

Lorsque les communes mirent sur pied des applications informatiques, les Archives reçurent - au mieux - les impressions papiers de documents faisant état de situations ou de décisions prises. Ce pouvait être des listings issus de bases de données, imprimés au moyen d'imprimantes à aiguille, faiblement encrés, sur des papiers de mauvaise qualité. Ces masses documentaires, lorsqu'elles ont été archivées, n'offriront à l'avenir qu'un faible intérêt analytique : sous réserve que les listings aient été produits selon un critère alphabétique, une recherche par nom peut seule être envisagée pour des contrôles biographiques. L'analyse sociologique des revenus n'est guère imaginable car des sondages fondés sur l'alphabet sont le plus souvent irrecevables et le coût de la numérisation avec passage d'une reconnaissance des caractères serait trop longue pour une recherche, sans compter les conditions pratique du scannage, les documents étant souvent reliés.

Dans les années '80, naissait l'ordinateur personnel et la gestion décentralisée des documents électroniques s'accéléra, posant du même coup un problème archivistique nouveau : comment repérer et évaluer les nouvelles applications élaborées par les différents services?

Alors que la production centralisée de la décennie précédente offrait une certaine garantie sur l'organisation logique des données, la micro-informatique a autorisé des développements tous azimuts, intéressants du point de vue opérationnels mais mal documentés: l'archiviste généralement ne sait rien des manières dont les contrôles et la gestion de l'information furent assurés dans les services.

Depuis les années '90, dans les grandes communes d'abord, la recherche d'une plus grande efficacité a conduit à la mise en place de bonnes conditions de transmission de l'information sur réseau et des groupes de travail modulables sont de plus en plus chargés de simplifier et rationaliser la production des documents. Les dossiers d'écriture qu'ils soient liés aux bâtiments, chantiers, cadastre, etc., sont et seront de plus en plus traités à travers l'informatique.

Aujourd'hui l'administration emprunte un chemin qui implique des modifications technologiques profondes. Celles-ci vont de pair avec des changements organisationnels (structures hiérarchiques comprises). Tout cela modifie les conditions de provenance, propriété et emplacement physique des données et l'archiviste doit se préparer à ces transformations.

Que ce soit pour des applications type bases de données du contrôle des habitants ou qu'il s'agisse de s'inscrire dans la logique des dossiers électroniques partagés, se pose ou se posera le problème de la conservation du document électronique lui-même plutôt que son tirage papier. Ce qui va être présenté ici ne saurait couvrir l'immensité d'un champ nouveau. On se contente de préciser quelques options *réalistes*, adaptées aux documents électroniques "classiques", c'est à dire non graphiques.

## Face à l'informatique : quelques « bonnes » questions

Au niveau communal, il serait tout à fait étonnant que les services désireux d'informatiser leurs prestations viennent consulter l'archiviste pour discuter de l'archivage de leurs données. Il convient donc pour l'archiviste de mener de front deux démarches :

1. Il lui faut recenser l'existant ou - ce qui peut arriver - ce qui a été créé sur informatique et qui a été abandonné. Cette tâche n'est pas aisée, car les créations "sauvages" ont fleuri et les collaborateurs de l'administration doivent jouer le jeu.

Tous les développements cependant ne sont pas d'un intérêt égal et il convient de hiérarchiser ses interventions, surtout lorsqu'on est seul ou presque. Le catalogue des systèmes informatiques a d'ailleurs un intérêt direct pour l'exécutif si celui-ci veut promouvoir une politique informatique cohérente au sein de sa commune. Cet aspect est très important, lorsque l'archiviste sollicitera sa hiérarchie pour justifier ce travail.

Un tel catalogue doit chercher à recueillir les informations suivantes:

Les éléments d'identification de la base devraient systématiquement contenir les éléments suivants :

1. Descriptif du système (objectifs)
2. Définition des structures d'enregistrement des données
3. Schéma conceptuel ou logique du système (relations entre les informations)
4. Créateur (service, bureau, nom de la personne)
5. Date de création
6. Position physique du système (support, accès partagé, monoposte...)
7. Logiciel utilisé, version
8. Type de mise à jour (maintien de l'historique, recouvrement des données anciennes)
9. Fréquence et modalités des mises à jour
10. Existence ou non de procédures de backup (sauvegarde)
11. Prise en compte de l'archivage (modalités retenues par le producteur)

Il s'agit donc de recevoir des documents fiables et authentiques : il faut que le document soit bien ce qu'il prétend être, qu'il conserve donc ses caractéristiques originelles et sa fiabilité première. Le contexte, la structure de données et le contenu sont autant d'éléments qui déterminent la qualité du document.

Archiver signifie donc stocker, documenter et donner les moyens de retrouver l'information. La documentation logique et technique des documents électroniques - les métadonnées - est primordiale. Face à l'existant, et surtout lorsque le producteur décide de laisser tomber son système informatique sans en reprendre le contenu, l'archiviste devra être très vigilant. Contrairement aux données sur papier qui peuvent être facilement conservées dans un local (un local sain bien sûr!) dont la clé est en main du chef de service, il faut agir très vite en matière informatique: l'évolution des systèmes, les connaissances qu'ils impliquent, mais aussi les supports, dont on doit souligner toujours la fragilité physique, risquent d'entraîner purement et simplement la perte de l'information, si on ne prend pas les mesures au plus vite. A ce moment là, des décisions devront être prises en accord avec le service producteur et l'exécutif pour garantir le maintien des informations sur un support dont la maintenance devrait incomber à l'archiviste. Si le stockage de l'information sur support électronique est judicieux - compte tenu de la masse des informations essentiellement -, on propose d'agir ainsi: il faut "exporter" les informations depuis le logiciel qui a servi à la gestion des informations pour générer un fichier en format Ascii (fichier plat). Un texte décrira le contenu de ce fichier soit l'ordre et la nature des informations qui y seront reportées, sans oublier les clés de lecture des codes qui auront été utilisés par le service producteur. Le stockage de ce fichier pourra être assuré sur des disques durs si leur taille le

permet avec copie sur bande; on pourra graver des CD-Rom comme sécurité. Il ne faut pas trop compter sur la qualité des CD-Rom gravés à partir de son poste de travail. Ces CD sensibles à la chaleur et à la lumière risquent fort de s'altérer s'ils sont régulièrement consultés.

En revanche, dans un contexte où les moyens humains et technologiques sont limités, si l'on sait que le service assure la maintenance du système *avec maintien de l'historique*, il vaut mieux ne pas chercher à effectuer un archivage des données par l'archiviste. Celui-ci devra plutôt réserver ses interventions dans les cas déjà vus d'abandon pur et simple du système ou de modification de la structure d'enregistrement des données. La procédure décrite sommairement ci-dessus pourra être reprise.

2. Lorsque l'archiviste a connaissance d'un nouveau projet pour lequel il n'a pas été consulté, il doit profiter de sa position pour rappeler quelques points cruciaux qui peuvent faciliter l'archivage futur des données.

Tout d'abord, il doit sensibiliser l'organisme qui tend à se pourvoir d'un système informatique, que l'intégrité, l'authenticité et le statut légal des documents émis est à valider et il doit faire valoir le besoin d'une documentation synthétisant précisément cette validation. Une telle prise en compte obligera le producteur à se poser des questions sur la gestion archivistique des dossiers et celle-ci pourra être plus facilement intégrée dès la phase de conception des systèmes.

### **Ne pas succomber à la tentation...**

Dans tous les cas, les conditions d'un archivage à long terme doivent être précisées et une attitude fortement sélective doit prévaloir. La baisse des coûts de stockage est un atout en faveur de l'archivage électronique mais il ne faut pas oublier les coûts de gestion afférents au maintien des données sur un support informatique. Des estimations américaines récentes ont montré que pour 1 mégaoctet de stockage, 1 dollar est consacré par année aux unités de disque contre 7 dollars à leur gestion !  
Préalablement à toute opération d'archivage électronique, il faut donc répondre aux questions d'évaluation suivantes :

Le volume d'informations du système justifie-t-il une conservation numérique ?

Existe-t-il un besoin permanent (fréquent ou non) de traiter les données à terme exigeant leur conservation numérique ?

Les informations ont-elles une valeur historique potentielle qui sera rehaussée par la conservation sur support électronique ?

Les informations du système ou document à archiver sont-elles un résumé des informations gérées par électronique dans un autre système ?

Ainsi, on préférera les sorties en clair ou le microfilmage pour la conservation d'informations statiques par nature, faisant l'objet d'annotations fréquentes ou constituées d'un nombre variable de types de documents. Sauf disposition contractuelle, les fichiers de traitement de texte, les tableurs par exemple, seront toujours traités sur la base d'un tirage papier avec désignation d'un exemplaire de référence. Les graphiques de présentation, les données statistiques graphiques seront également conservées sur papier. En revanche, il faut privilégier un archivage numérique chaque fois que l'information nécessitera le recours à l'électronique pour une analyse ou manipulation ultérieure.

## Orientations pour sauvegarder les fichiers

On veillera à un entreposage inerte (à l'abri des champs magnétiques) et en position verticale. Eviter les chocs thermiques (température stable de 17 à 20 degrés, humidité relative de 35 à 45%). Le plus grand risque est la perte des signaux magnétiques. La puissance des signaux se dégrade lentement mais sûrement sur les supports en oxyde de fer. Il faut donc prévoir des copies régulières de l'information.

Les supports en dioxyde de chrome sont utilisés massivement depuis 1985. Ils sont beaucoup plus stables en l'absence de toute perturbation du support (rebobinage, tension).

Dans les deux cas, le substrat des supports absorbe l'eau. Trop d'humidité ou trop de chocs hygrométriques se traduisent par la déformation voire le détachement du revêtement magnétique qui est sur le substrat. Les signaux deviennent alors illisibles.

Pour les disques souples, il n'existe aucune norme pour la conservation à long terme, on s'en méfiera donc. De même, il n'existe pas non plus de norme pour les disques optiques et les risques d'obsolescence sont élevés à moins que des budgets ad hoc soient prévus pour le transfert des données.

disque dur à accès immédiat	prévoir des sauvegardes sur bandes ou cartouches sur 2 sites
disquettes (3.5 pouces)	à bannir au delà d'une année. Vérifier leur validité au moment de la copie
bandes 1/4 pouce et cartouches à balayage hélicoïdal	uniquement pour des sauvegardes de reprise après destruction de systèmes. Les formats d'enregistrement n'étant pas normalisés, ne pas utiliser pour un archivage long terme
bandes en oxyde de fer(streamer)	veiller à les tester régulièrement (tests d'erreurs permanentes). Rebobiner la bande sous tension contrôlée.
bandes en dioxyde de chrome	ne pas lire, retendre ou nettoyer ces bandes. Tester la lisibilité à échéances régulières.

Pour conclure, il semble utile de rappeler que la conservation des technologies liées à la production des documents électroniques, dans une sorte de musée vivant de l'électronique, n'est pas une solution envisageable pour les archives communales. Le coût financier et humain serait rapidement gigantesque du fait de l'obsolescence technologique rapide des systèmes et de la perte des savoirs qui s'y rapportent. A la question : "qui saura encore utiliser un IBM 3090 après 20 ans d'oubli?", nous ne doutons pas de la réponse: personne.