

Le patrimoine culturel numérique

Le patrimoine archivistique numérique : les défis de sa conservation à très long terme. Aspects stratégiques et opérationnels autour de la solution des Archives d'Etat de Genève.

Anouk Dunant Gonzenbach

1. Les Archives d'Etat de Genève (AEG)

La Suisse est constituée de 26 cantons qui sont autant d'Etats souverains ayant délégué certaines compétences à un Etat fédéral dans le cadre de la Constitution. Chaque canton est pourvu d'une institution d'archives cantonales et de sa propre législation relative aux archives. Les pratiques, les instruments de travail ou les logiciels d'archives ne sont pas centralisés; chaque canton développe donc ses propres outils, notamment ses outils administratifs.

Les archives cantonales conservent des documents de l'administration du canton, du moyen-âge jusqu'à nos jours, ainsi que des documents de provenance privée, soit des archives de familles, de sociétés, d'associations ou d'entreprises. L'ensemble des fonds conservés aux AEG représente environ 29 kilomètres linéaires de documents, depuis le plus ancien, daté de 1099, jusqu'à aujourd'hui. La mission des AEG est de conserver les documents produits par l'Etat de Genève et évalués comme historiques sur le très long terme.

2. La nécessité de mise en œuvre de l'archivage électronique

L'information est désormais créée, enregistrée et transmise électroniquement. L'échange de données et de documents entre l'Etat et les citoyens s'est transformé : les processus se dématérialisent ou sont déjà dématérialisés à travers les projets de cyberadministration. Pourtant la mission des archivistes, elle, reste la même et les documents électroniques ayant une valeur juridique, politique, patrimoniale, économique ou historique doivent être conservés sur le très long terme. Il est ainsi par exemple indispensable d'archiver le registre foncier ou le cadastre (la série conservée aux AEG débute en 1689). Or le cadastre est désormais entièrement dématérialisé dans le système d'information du territoire genevois et il n'est plus question de l'imprimer.

Il est nécessaire d'apporter ici une précision: mon propos concerne les documents et données nés-numériques, qui n'existent pas et n'existeront probablement pas sous forme papier, et non pas les images numérisées. Lors de projets de numérisation d'archives, les originaux sont conservés. Ils bénéficient d'un programme de numérisation dans l'objectif de permettre leur diffusion, qui doit être la plus large possible. La mise en ligne d'images numérisées répond à une forte demande des chercheurs d'accéder à distance aux sources. Le deuxième objectif de la numérisation de documents est de préserver des originaux très consultés, donc très manipulés.

3. Le préalable: la politique d'archivage électronique

Il n'est pas possible de mettre en œuvre un projet puis une solution d'archivage électronique pérenne sans soutien et approbation au plus haut niveau hiérarchique, dans notre cas au niveau du Conseil d'Etat de Genève, le gouvernement exécutif du canton. En

effet, la gestion de l'archivage électronique implique des ressources humaines et financières dont il faut être conscient. En réalité, l'archivage n'est pas une préoccupation essentielle. Pourtant la prise en compte de la problématique de pérennisation des informations numériques est de la responsabilité des dirigeants; négliger cette question aura pour conséquence la perte d'informations essentielles au fonctionnement de l'Etat, dans un terme situé entre cinq et quinze ans. En revanche, l'information adéquatement conservée et pérennisée sera facilement accessible à l'administration, aux citoyens et aux chercheurs, et cela conformément à la législation.

Avant de mettre en place une solution d'archivage électronique, il est nécessaire de préciser les objectifs et le périmètre de la problématique. Nous avons donc tout d'abord rédigé un document intitulé "politique d'archivage à long terme des documents électroniques", très général, sans processus, un cadre pour l'implémentation de l'archivage électronique. Ce cadre pose l'environnement juridique, le contexte de l'archivage électronique, sa définition, les points forts et les partenaires en présence. Le Conseil d'Etat a validé ce document, ce qui a permis de mettre en place la solution d'archivage pérenne.

4. La communication

Une communication adéquate est essentielle pour sensibiliser tous les partenaires et obtenir les ressources nécessaires. En effet, la problématique de l'archivage n'est une priorité pour personne. La communication est l'un des facteurs clés pour mener un projet d'archivage électronique et obtenir son financement. Il appartient aux archivistes de s'adapter, de dépoussiérer leur discours et de le traduire en des termes compréhensibles par les professionnels des systèmes d'information. Nous devons "nous vendre" tout en restant attentifs au vocabulaire, car beaucoup de mots n'ont pas la même signification pour un archiviste, un informaticien ou un juriste. Pour un informaticien par exemple, le long terme signifie une quinzaine d'années au maximum. Si l'on ne fait pas attention à ces notions perçues différemment entre les professions, la discussion aura lieu sur un malentendu. Nous avons ainsi mis en place une stratégie de communication, utilisée dans les nombreuses présentations que nous avons faites aux informaticiens, aux juristes et aux décideurs.

Je crois qu'il est également important de préciser ici que nous sommes vraiment arrivés à intervenir en tant que partenaires dès le moment où l'on s'est attaqué aux problématiques qui touchent de près les informaticiens, les juristes et les directeurs de service, à savoir la gestion des documents électroniques au sein de l'administration, et notamment la question difficile et cruciale de l'authenticité du document électronique, pour laquelle nous pouvons offrir notre savoir-faire.

5. La solution Gal@atae d'archivage électronique à long terme

Les AEG ont mené entre 2011 et 2013 le projet Gal@tae (**Genève** archivage à long terme d'**archives électroniques**). Le choix de ce nom se réfère à Galatée, la statue dont le sculpteur Pygmalion tomba amoureux et à laquelle Aphrodite donna la vie. Ce nom rappelle que des archives "vivantes" se trouvent derrière l'aspect technique de l'informatique. Galatée, qui est également le nom d'un satellite de Neptune, évoque la possibilité de conserver les archives électroniques au loin, hors des murs et du territoire exigü de Genève.

Il s'agit d'archiver des documents électroniques enregistrés dans des formats soigneusement sélectionnés (PDF-A par exemple), pourvus de métadonnées techniques et descriptives. Ces documents sont conservés dans des répertoires appelés "paquets d'archives", ou SIP. La stratégie est de migrer régulièrement ces paquets d'archives pour éviter l'obsolescence de leurs formats.

Du fait de la nature des documents électroniques, leur archivage doit être prévu dès le début de la production documentaire, à savoir dès la création des systèmes d'information. Il faut les pourvoir de métadonnées et les enregistrer dans les formats adéquats dès leur création, car il est techniquement beaucoup plus compliqué de faire toutes les conversions en bout de chaîne. Les documents numériques archivés doivent être intègres, fiables, authentiques et exploitables sur le très long terme.

Le standard en matière d'archivage électronique à long terme est la norme OAIS (système ouvert d'archivage de l'information), enregistrée ISO sous la référence 14721:2002. Il s'agit d'un modèle conceptuel destiné à la gestion, à l'archivage et à la préservation à long terme des documents électroniques.

Dans ce modèle, on trouve tout d'abord évidemment les producteurs d'archives, dans notre cas les services de l'administration. Comme producteur de documents électroniques, nous avons choisi pour le projet pilote le service des autorisations de construire. Dans le cadre des projets de cyberadministration, les architectes pourront déposer en ligne leurs demandes d'autorisation de construire et leurs plans. Les AEG conservent les dossiers d'autorisation de construire sous forme papier depuis 1848 et leur consultation est très fréquente. Il était donc indispensable de prévoir l'archivage des plans sous forme électronique pour éviter une rupture dans la continuité de la série. Il était également important de profiter de l'opportunité présentée par le projet de dématérialisation de ces dossiers pour tester leur archivage.

Trois possibilités s'offraient pour l'hébergement des archives électroniques: choisir un hébergeur privé, construire une plate-forme de pérennisation au sein de l'Etat de Genève ou utiliser une plate-forme existante institutionnelle, en l'occurrence la plate-forme des Archives fédérales suisses (AFS). Nous avons choisi de collaborer avec les AFS, dont la plate-forme offre plusieurs avantages : en conformité à la norme OAIS, elle est hébergée par une administration publique, donc pérenne. Cette plate-forme a été fortement dotée en moyens financiers par la Confédération et met à disposition toute l'infrastructure technique, assurant l'hébergement sur trois sites géographiques. Le contrôle des données, de leur accès et de leur protection (au sens de la législation genevoise sur la protection des données) se fait par l'administration genevoise, à savoir par les AEG.

Les utilisateurs sont bien entendu le public (chercheurs, historiens, etc.) mais également les services de l'administration qui auront recours à des documents numériques ne se trouvant plus dans leur système d'information.

Dans le cas du papier, les documents conservés à long terme sont traditionnellement conservés dans des boîtes d'archives en carton. Par analogie, en archivage électronique, les documents électroniques sont conservés dans une boîte virtuelle, un paquet d'information (SIP). Un paquet d'information est constitué des données primaires (les documents électroniques que l'on veut conserver) et des métadonnées indispensables à leur conservation.

L'archivage s'effectue selon le processus suivant: le service administratif sélectionne les documents à conserver et utilise une application développée à cet effet pour les placer dans des boîtes virtuelles (les SIP). En plus des SIP, l'application génère un fichier XML contenant les métadonnées de description qui sera importé dans la base de données des archives. Ces SIP sont ensuite versés aux AEG, où sont soigneusement contrôlés leur structure, leurs métadonnées, leur format et leur nommage.

Une fois validés par les AEG, les SIP sont transférés aux Archives fédérales par un processus de transfert sécurisé. Ils y subissent des tests pour assurer leur conformité. Les Archives fédérales ajoutent également un certain nombre de métadonnées techniques propres à leur gestion de la plateforme de pérennisation et par un processus d'ingest versent les SIP sur leur plate-forme de pérennisation.

Le chercheur qui voudra consulter des données numériques archivées effectuera une recherche sur la version en ligne de la base de données des AEG. A terme, les documents consultables seront disponibles en ligne directement sous format numérique. Il y aura donc une version de diffusion, en ligne, et la version « authentique », dans le dépôt numérique. Les documents qui contiennent des données personnelles ou qui sont encore protégés ne seront évidemment pas consultables.

La solution Gal@tae est opérationnelle depuis janvier 2014.

6. Conclusion

Il n'y a pas de système standard d'archivage électronique proposé et utilisé par les différentes institutions cantonales en Suisse. Les cantons développent donc leurs propres solutions. Les avantages de la solution « Archives fédérales » nous ont très vite apparus évidents. Nous sommes en effet convaincus par la mutualisation d'une plate-forme de pérennisation, qui permet une réduction des coûts de maintenance et de migration. J'insiste sur le fait que le contrôle de la consultation, et donc des durées de protection des documents, se fait aux AEG; la législation cantonale est donc appliquée. Ce modèle de collaboration entre les cantons et les Archives fédérales convient à notre sens parfaitement.

Un des éléments fondamentaux d'une telle solution est le contrat entre les deux parties, à savoir une convention et le Service Level Agreement. Ces documents règlent les modalités de l'archivage, les garanties techniques et tous les aspects liés à la sécurité et à la confidentialité des données entre les deux institutions.

Pour obtenir des données archivables, les archivistes doivent agir en amont de leur terrain d'action habituel et sortir de leur institution. Ils doivent être présents déjà dans la conception des systèmes d'information. En conséquence, la profession d'archiviste évolue, ou se métamorphose. L'archiviste doit acquérir de nouvelles compétences et remonter la chaîne de la production documentaire.