

## - Le blog de Marie-Anne Chabin

---

 [marieannechabin.fr/archiver-et-apres/3-du-papier-au-numerique/](http://marieannechabin.fr/archiver-et-apres/3-du-papier-au-numerique/)

### *Archiver, après ?* Marie-Anne Chabin, Djakarta, 2007

---

#### Chapitre 3 – Du papier au numérique

---



*Il vaut mieux penser le changement que changer le pansement*

Pierre Dac

Si le verbe archiver a du succès aujourd'hui, c'est indéniablement à l'informatique qu'il le doit.

L'archivage est lié à la notion de masse, indissociable d'un certain volume d'information. Si on a trois photos ou dix courriers qui datent de l'année dernière ou de l'année d'avant, on n'a pas de problème d'archivage. Les documents sont là, à portée de main et de mémoire, et ils y resteront si rien ne vient s'entremettre pour les cacher à la vue et aux préoccupations quotidiennes.

Mais l'informatique accroît considérablement la masse d'informations, toujours plus difficile à maîtriser. Après l'euphorie de la création numérique sans contrainte, vient la crainte de la disparition de la mémoire. Les jeunes informations, toujours plus nombreuses et toujours plus offensives, poussent les vieilles vers le fond de la mémoire, vers l'oubli. La nouvelle vague de données balaie la précédente, menace d'emporter sur son passage tout ce qui a le malheur de dater un peu, l'accessoire, dont la perte sera peu dommageable et qui même donnera un peu d'air, mais aussi l'essentiel...

Le besoin d'archiver se fait alors plus prégnant. Mais peut-on comparer l'archivage numérique à l'archivage traditionnel des documents papier ou des films ?

#### **L'inforrhée galopante**

---

L'expression « société de l'information » dit bien ce qu'elle veut dire : l'information est partout, à tous les étages de la société et de plus en plus envahissante.

L'information est largement numérique et couvre toutes les formes de l'expression et de l'activité humaine. Elle représente aussi bien du texte ou des chiffres, héritiers des écrits traditionnels, que des images, du son ou des signaux, héritiers des supports analogiques issus des outils de capture du réel apparus au XIXe siècle : photographie, disque, films, vidéo.

On admet que 90 à 95 % de l'information est aujourd'hui produite et diffusée à l'aide d'outils informatiques. Mais une proportion non négligeable des textes numériques est ensuite matérialisée en papier, soit sous forme de livres ou de brochures publicitaires, soit sous forme de rapports, procédures, courriers ou de simples documents de travail. L'inflation de la production d'information est telle que, même si l'habitude d'imprimer systématiquement textes et photos se perd peu à peu, le volume de papier sorti des ordinateurs augmente encore en valeur absolue. Le zéro papier qu'on nous promettait dans les années 1980 n'est toujours pas au rendez-vous. Il faut du temps pour que les usagers s'approprient les arguments de la culture de l'écran, du développement durable ou de la simple maîtrise du gaspillage.

Comment décrire cette réalité de 0 et de 1 qui nous entoure et s'installe inexorablement dans notre vie ? L'analyse quantitative est instructive mais l'approche qualitative est plus intéressante : données scientifiques, administratives, commerciales ou culturelles, nombre de documents ou objets de données, évolution du poids moyen de ces objets, redondances.

De même que le papier se compte en pages, en mètres linéaires puis en kilomètres linéaires de rayonnages, de même que l'image animée et le son se comptent en heures, l'information numérique, pour sa part, se compte en octets (en anglais *byte*), un octet étant l'unité de base pour le codage d'un caractère d'écriture, une couleur ou d'un signal, à l'aide de 0 et de 1. Avec les progrès technologiques, le codage de l'information requiert toujours plus d'octets et génère régulièrement des nouvelles unités de mesure forgées selon la progression arithmétique des puissances de 10 : le millier d'octets ou kilooctet (Ko ou  $10^3$ ) suffit encore pour transcrire quelques phrases ou une imagerie ; le mégaoctet ou millier de kilos (Mo ou  $10^6$ ) qui correspond aujourd'hui à un rapport de 100 pages illustré ou à une belle photographie ; le gigaoctet (Go ou  $10^9$ ) qui équivaut au poids de la boîte de messagerie électronique d'une secrétaire de direction ou d'un examen médical. L'heure est aux téraoctets (To) mais ils seront bientôt dépassés par les péta (Po), puis par les exas (Eo) et ainsi de suite au gré des préfixes simili-grecs qui servent cette cause (zetta, yotta,...).

Les plus gros producteurs de données, en volumes, sont les centres scientifiques avec la capture de données spatiales ou géophysiques par satellites ou les traitements mathématiques à grande échelle. Les archives numériques du Centre national d'Études

Spatiales (CNES) pèsent aujourd'hui près de 500 To de données regroupées en plusieurs millions de fichiers. Le Commissariat à l'Énergie Atomique (CEA) a mis en service en 2006 le calculateur le plus puissant d'Europe dont la capacité totale atteint sept Po...

L'imagerie médicale numérique se développe à la même allure. Non seulement on pratique davantage d'examens, mais le nombre et le poids (c'est-à-dire la précision) des clichés augmentent considérablement. Un hôpital qui stocke aujourd'hui quelques To d'images en aura aisément cinq ou dix fois plus d'ici dix ans. Toutes ces données, plus facilement accessibles que des chemises cartonnées empilées sur un bureau ou qu'il faut « remonter » de la salle d'archives, permettent de soigner mieux.

A côté des données de recherche et de traitement, un deuxième ensemble est constitué par des documents de communication institutionnelle ou commerciale : livres, radio, télévision, cinéma, publicité, qui se présentent de plus en plus sous forme numérique. Là aussi se conjuguent la multiplication des acteurs qui produisent de l'information, notamment sur Internet, et l'augmentation des volumes due au progrès technologique qui permet une plus grande définition et une animation de l'image. *Internet Archive*, initiative privée américaine lancée en 1996 pour sauvegarder la mémoire d'Internet stocke aujourd'hui plus de 100 téraoctets de pages web, tandis que pour la seule année 2004, la Bibliothèque nationale de France (BnF) en a capturé 30 To.

Les autres données, qualifiées par les informaticiens de « non structurées » ou « semi-structurées » (par opposition aux données bien carrées des grandes bases de données, faites de chiffres ou de mots bien organisés) forment un troisième ensemble. Il s'agit essentiellement de textes, plus ou moins illustrés : d'une part, les rapports et études, notes et comptes rendus, courriers électroniques, présentations pour conférences, tableaux, etc. qui prennent la succession de l'écrit papier classique ; d'autre part, les pages personnelles d'Internet, les forums de discussions et les blogs dont le genre a éclos sous la baguette magique de la technologie.

Et l'information grossit, grossit. On recopie, on colle, on versionne, on corrige, on crée un lien, on re-recopie, on diffuse, on complète, on rediffuse, on reformate, on re-colle, on modifie, on compile, on re-re-recopie, etc. Les serveurs bureautiques arrivent à saturation mais les systèmes, toujours plus performants, permettent d'aller de l'avant. Alors, on va de l'avant. Les outils sont « performants » ; ils savent garder la trace de la moindre modification apportée à un document. Une entreprise a ainsi dénombré jusqu'à 562 versions de la même note, c'est formidable. Quant aux messages électroniques, si l'on en croit les prévisions, ils augmentent de 25, 50 voire 100% par an : 10 milliards de courriels échangés en 2001, 35 milliards en 2005, 60 milliards en 2006...

La comparaison avec la progression géométrique des nénuphars sur l'étang est dépassée...

Dans un sens, tout cela est normal. Le progrès technologique se poursuit. La loi informatique de Gordon Moore dit que la puissance des machines double tous les dix-huit mois, avec des variantes selon que l'on parle de la capacité de stockage ou de la

capacité des réseaux, tandis que le coût matériel décroît.

D'un autre point de vue, si les pétaoctets de données scientifiques donnent le vertige, ce flux grossissant et bouillonnant de données diverses et variées qui circule dans les réseaux, cette véritable inforrhée galopante, donne le tournis.

*Quousque tandem ? Jusqu'où ?*

**On se réjouissait de voir l'informatique faire disparaître toute la paperasse mais si on l'a troquée pour de l'électronasse, où est le progrès ?**

Que faire de toutes ces masses ? Ou bien elles sont utiles et il faut les archiver et en payer le prix, ou bien elles ne présentent pas d'intérêt et pourquoi s'en encombrer ?

## **Quelques idées reçues**

---

La technologie numérique représente plus qu'une substitution de support de l'information. C'est une véritable révolution dans la conception même de l'organisation de l'écrit. Le numérique induit des changements radicaux des modes de production et de diffusion, dans l'économie, dans le droit et la culture, et bien sûr dans les pratiques d'archivage et de conservation.

A cet égard, il n'est pas simple d'appréhender toutes les composantes de cette problématique avec le recul nécessaire, d'autant plus que la technologie évolue sans cesse. C'est pourquoi on rencontre ici et là, même chez des décideurs, des responsables informatiques ou des archivistes, des opinions, ou plutôt des idées reçues, quelquefois contradictoires, trop rapidement transposées des principes d'organisation du monde papier.

D'aucuns pensent par exemple que la technologie va mettre au point un support numérique fiable et pérenne, une sorte de parchemin numérique, que l'on pourra ranger en toute confiance sur des rayonnages, à côté des livres et des registres de l'ancien temps, et que, dans cent ans, on pourra le relire sans difficultés. Alors que, comme l'affirme très justement le groupe de recherche international sur l'archivage numérique InterPARES : « Conserver un document électronique est à proprement parler impossible ; seule la capacité de le reproduire peut être préservée ». Le support numérique ne dure pas. Il mute. Il migre.

D'autres estiment que le numérique ne convient pas pour l'archivage. Pour deux raisons principalement : ce n'est pas fiable et c'est trop cher. Certes, les supports numériques n'ont pas la résistance intrinsèque des supports traditionnels que sont le microfilm ou le papier chiffon mais la question ne doit pas être posée en termes manichéens comme si la seule alternative se situait entre le bien-papier et le mal-numérique ou, pour d'autres, entre le bien-numérique et le mal-papier.

Que seraient devenues les premières bandes vidéo de la télévision, de la fin des années 1950, sans les supports numériques ? Rien, car ils auraient tout bonnement disparus, non pas tant à cause de l'effacement des bandes qu'en raison de la disparition totale des appareils capables de les lire. En transférant les données sur un support numérique pendant qu'il en était encore temps, l'Institut national de l'Audiovisuel (INA) a sauvé des centaines d'heures d'émissions et nous sommes assurés de pouvoir les visionner dans les décennies à venir. La remarque n'est pas valable que pour la vidéo ou les disques vinyles : les journaux imprimés pendant la Seconde Guerre mondiale ou les pelures administratives de cette époque courent le même risque de détérioration et ne peuvent regarder le numérique que comme un sauveur.

Quant au coût, il convient de nuancer, nous y reviendrons dans le chapitre suivant. Suggérons simplement un petit calcul : un kilomètre linéaire d'archives papier nécessite une surface de stockage de 90 à 170 m<sup>2</sup> selon le type de rayonnages ; l'équivalent numérique d'un kilomètre linéaire de papier (texte) est 100 Go (que l'on peut stocker sur une seule bande) ; et le coût du mètre carré est de 2000 € le m<sup>2</sup>... D'où il ne faudrait pas conclure non plus qu'il faut se dépêcher de tout numériser.

Une opinion, assez fallacieuse, s'est répandue également selon laquelle l'utilisation généralisée de l'informatique va conduire progressivement à une disparition de l'archivage. Avec les téléprocédures, plus de dossiers papier, la machine s'occupe de tout. Sans doute, sans doute, mais comment seront gérées les données dans les années à venir ? Comment les retrouver après des années d'accumulation d'autres données ? A qui seront-elles accessibles ? Selon quels critères seront-elles détruites ou conservées ? Éluder ces questions revient à confondre naïvement archivage avec stockage ou sauvegarde. Pour la mémoire, la technologie est fondamentale mais ne suffit pas : science sans conscience...

Même si les choses évoluent rapidement sous l'effet de la jurisprudence américaine qui met en avant la valeur juridique des messages électroniques et ce qu'il en coûte à certains dirigeants négligents de les détruire, l'idée que les courriels n'ont pas vocation à être archivés est encore bien ancrée.

La messagerie électronique suscite des pratiques très variées, plus sans doute que toute autre moyen de communication. Il y a ceux qui les impriment, ceux qui les empilent dans leur boîte de messagerie, ceux qui les suppriment au fur et à mesure, ceux qui les refilent au voisin. Il y a ceux qui utilisent la messagerie pour dire « Bonjour vieux, ça va ? » et ceux qui utilisent la messagerie pour dire « OK, je vous fais une ristourne de 20% sur ce contrat de deux millions d'euros ! ». Il y a ceux qui s'expriment de manière explicite, comme dans un courrier classique, et ceux qui sont allusifs, elliptiques, instantanés.

Au-delà des pratiques individuelles de gestion de l'information dans son environnement personnel ou professionnel, la question n'est pas de savoir si, par principe, il faut archiver les courriels ou s'il ne faut pas les archiver ; la question est de savoir ce que dit ce

message, ce qu'il représente, ce qu'il sous-entend, bref ce qu'il induit pour celui qui l'envoie comme pour celui qui le reçoit voire pour des tiers qui seraient ultérieurement amenés à s'y référer.

Au nombre des idées reçues, on pourrait citer également la croyance selon laquelle en appuyant sur la touche « supprimer » de l'ordinateur on a effectivement détruit le fichier, alors qu'il est toujours là, caché dans la mémoire informatique où un expert saura le dénicher. Ou encore la conviction qu'il est préférable d'imprimer un document numérique, signé électroniquement, parce que « c'est le papier qui fait foi », alors que le fichier numérique constitue l'original au regard de la loi.

Les habitudes évoluent toutefois.

## **Ce qui change vraiment pour l'archivage**

---

Le processus de passage du papier à l'électronique dans la gestion courante de l'information n'est pas achevé. C'est un processus qui s'étire entre les années 1980 et les années 2010. Mais on voit bien déjà, outre l'explosion des volumes, les caractéristiques fondamentales du changement. Il y en a au moins quatre.

Tout d'abord, c'est la dimension immatérielle qui change la donne. L'archive numérique n'est pas un objet visible et palpable comme l'est un registre ou un film 16 mm, un objet que l'on peut feuilleter, dérouler simplement avec ses deux mains, dont on peut voir le contenu, même si on ne le déchiffre pas.

L'archive numérique n'a pas d'odeur, de cette odeur de matière qui se bonifie avec les ans et à laquelle l'archiviste est attaché, bien qu'elle ne soit pas objectivement agréable, de même que le mineur est attaché à l'odeur du charbon. Quand on aura dit que l'archive numérique n'émet pas de bruit comparable à la page que l'on froisse ou au film que l'on déroule, et qu'elle n'a pas non plus de goût, ce goût de l'archive vanté par Arlette Farge[1], on constatera que la frustration sensorielle est complète.

Eh oui, l'archive numérique est un objet virtuel, indirect, conditionnel, dont l'appréhension est soumise au truchement d'un matériel, lui-même dépendant d'un système et d'un réseau technologique qui en permettront, à certaines conditions, la visualisation.

Mais, on peut toujours croire que, pour compenser, l'archive numérique est accessible à un sixième sens...

La seconde caractéristique de l'archivage numérique touche à l'accès à la matière archivée. Les réseaux abolissent les distances. Si la finalité de l'archivage est de pouvoir accéder ultérieurement à l'information (plus ou moins rapidement selon les besoins), avec le numérique, la contrainte de devoir archiver au plus près de l'utilisateur est levée. La mise en commun de documents scannés est depuis longtemps déjà pratiquée au travers des outils de gestion électronique de documents (GED) mais l'argument vaut aussi pour la conservation dans le temps. Il est acquis qu'il est inutile d'archiver le Journal officiel électronique dans les 36 000 communes de France, comme on le faisait

naguère pour l'édition imprimée ; une fois suffit puisqu'on peut avoir 36 000 accès pour un seul archivage. De la même façon, il est inutile d'archiver à la fois dans le département et en centrale tous les dossiers traités sous forme numérique par les préfets et transmis aux ministères ; il suffit de les archiver une fois au nom de l'État et de partager l'information entre Paris et la province.

On pourrait poursuivre le raisonnement pour l'état civil en considérant qu'un seul grand registre, ou plus exactement une seule grande base de données, pourrait enregistrer les naissances, mariages et décès sur l'ensemble du territoire national. De ce point de vue, l'archivage électronique va bousculer l'organisation administrative.

La troisième spécificité est la plus cruciale. Il s'agit de la faible espérance de vie de l'objet numérique. Le numérique est fragile et fugace, ce qui ne le prédispose pas à la vie d'archive qui, par nature, se veut pérenne. Les 0 et les 1 sont toujours là mais les intermédiaires qui en conditionnent le décodage et la lisibilité ne sont pas viables comme l'est un support d'écriture d'origine animale ou végétale.

Les exemples ne sont pas rares de données des années 1980 ou 1990, stockées sur disquettes dans des formats abandonnés ou dans des systèmes non évolutifs et qui ont été perdues. Sans doute ne faut-il y voir qu'un incident de parcours au regard de l'Histoire des techniques. La technologie apporte aujourd'hui des réponses satisfaisantes à ce handicap de l'objet numérique archivé. Le processus de pérennisation, au travers de matériaux et de formules mathématiques bien choisis et mis sous observation permanente, comme des plantes que l'on greffe, que l'on arrose, que l'on transplante avec soin, le processus de pérennisation numérique donc, est de nature à rivaliser avec les supports traditionnels.

Ainsi, ce n'est pas tant le danger d'obsolescence qui est à craindre dans cette affaire, c'est le risque de ne pas en prendre la mesure et de ne pas l'anticiper. Par rapport à un archivage papier qui prenait son temps, qui courait volontiers derrière le train de l'accumulation de dossiers, l'archivage numérique ou électronique (les mots sont ici équivalents) ne tolère pas d'attendre pour voir : il exige de conduire la locomotive !

Les possibilités d'automatiser l'archivage, entendez la capture et le contrôle des données, constituent la quatrième caractéristique de l'archivage numérique. C'est la plus intéressante sur le plan archivistique.

En effet, la mise en archive, c'est-à-dire le passage de l'environnement de création au lieu de conservation, n'est plus nécessairement une opération humaine. Il peut être en grande partie automatisé et fluidifié, à condition bien sûr que l'identification de l'objet à archiver et le traitement à lui appliquer aient préalablement été appris à la machine. Ainsi la qualification de l'objet archivé, qui fait appel à la compétence de l'archiviste, intervient dès la production voire dès la conception de l'information et non *a posteriori* comme cela se pratique communément.

Par ailleurs, la technologie numérique permet facilement la duplication ou la réplication des données, extrêmement lourde et coûteuse avec d'autres supports. La question de la sécurité des archives ne se pose donc plus dans les mêmes termes : la copie de sécurité n'est plus une opération manuelle à la charge d'une personne mais un automatisme ; la confection de microfilms de conservation (le microfilm reste dans certains cas un support valable) ne passe plus par la prise de vue du papier par un opérateur mais par un programme informatique.

## **Ce qui est toujours vrai**

---

Cependant, tout n'est pas bouleversé dans l'archivage électronique car la technique n'est que le contenant de l'information, laquelle reste produite par l'homme et pour l'homme, de sorte que les concepts fondamentaux de la gestion de l'information demeurent.

L'archive, numérique ou pas, reste composée d'un message qui a du sens pour celui qui l'émet et pour celui qui le reçoit, matérialisé par un ou plusieurs éléments physiques qui, tout à la fois, supportent le message au travers d'un code d'écriture et en assurent la transmission à son destinataire. Ce peut être une matière travaillée comme l'argile ou la pierre gravée, ou une matière modifiée par l'ajout d'une autre matière, comme le papier et l'encre, ou encore toute une chaîne d'éléments comme peuvent l'être un disque gravé + une baie de stockage + un logiciel + un ordinateur + un système d'exploitation + une prise de courant.

Finalement, la permanence de l'existence d'un support pour toute information archivée s'avère plus importante que la modification de la matérialité de ce support. Et le support d'archivage doit être géré, dans sa globalité et dans sa double valeur de support d'enregistrement et de support de transmission. C'est pourquoi, graver ses fichiers sur des cédéroms et les ranger dans un tiroir n'est pas archiver ; il manque la dimension de transmission. Le cédérom est tout à fait pratique lorsqu'on recherche, l'année suivante ou l'année d'après, le fichier qu'on se rappelle y avoir mis. Ça ne vaut plus rien au bout de quelques années, quand on a changé d'ordinateur et que les éditeurs ont publié deux nouvelles versions du format de codage, sans compter qu'on a peut-être oublié, avant de quitter le service, de prévenir les collègues de ce qu'ils pourraient y trouver, et que de toutes façons, on aurait été bien en peine de le faire vu qu'on a soi-même totalement oublié ce qu'on avait enregistré là-dedans !

Il convient d'insister sur le fait que le support numérique n'est pas un morceau de matière mais une chaîne matérielle et logicielle. Si cette chaîne n'est pas maintenue en état de marche, si un maillon de la chaîne défaille, le support est tronqué, inefficace, incapable de transmettre l'information, comme le serait un feuillet écrit à l'encre qui aurait séjourné dans la rivière ou une tablette d'argile qui serait passée sous un semi-remorque.

La notion de support, avec sa nécessaire transposition, reste donc pertinente dans l'archivage numérique.



La notion de fonds d'archives perdure elle aussi, même si on n'en parle pas et que des pratiques disparates masquent cette réalité juridique.

Un fonds d'archives est constitué de l'ensemble des informations de traçabilité et de mémoire créées ou reçues par une personne, une entreprise ou une institution au cours de son existence. Cette définition se recoupe avec celles de patrimoine informationnel et de mémoire d'entreprise. L'information archivée appartient à l'entité juridique au nom de laquelle et au sein de laquelle elle a été constituée, non à son rédacteur ou à son signataire, même si celui y a sa part de responsabilité. Ceci a peu d'importance quand la conservation se déroule sans incident, ce qui arrive dans la majorité des cas, fort heureusement. Mais si quelque chose se passe mal (divulgaration d'informations confidentielles, conservation abusive de données nominatives ou destruction induite d'informations, perte d'un contrat ou d'un titre de propriété intellectuelle, etc.), ce sont bien les plus hauts responsables de l'entreprise, de l'organisme ou de l'association qui sont comptables de ces erreurs devant les autorités civiles et judiciaires.

L'archivage ne peut donc s'envisager valablement qu'au niveau du fonds, au niveau global de la structure juridiquement agissante, au niveau du propriétaire ultime de l'information, de façon à assurer une cohérence aux règles d'archivage et de conservation, quel que soit le support. Cependant, dans l'environnement électronique bien plus que dans le monde papier, on observe deux tendances incongrues qui vont à l'encontre de cette logique : la personnalisation et la segmentation.

On ne sait pourquoi, sans doute pour séduire les utilisateurs ou simplement par ignorance, les logiciels informatiques poussent à l'appropriation par l'individu de l'information qu'il crée, ce qui dans le monde du travail aboutit à des aberrations : les documents que je produis dans le cadre de mon travail sont « Mes documents » et les courriers que je gère pour mon service sont « Mes courriers ». Le droit français est venu tout d'abord conforter cette attitude en considérant le message électronique échangé sur le lieu de travail comme une correspondance privée d'où on peut déduire que le message envoyé par un commercial de la société X à un client Y est une correspondance privée et que, si le commercial tombe malade et ne peut suivre l'affaire, son collègue ou son patron n'ont *a priori* pas le droit de lire ni de récupérer ce message. Il en découle que la correspondance professionnelle d'un collaborateur, directeur ou employé, ne peut pas être archivée au nom de l'entreprise ou au mieux que, si elle est archivée, ce sera à titre personnel et que personne d'autre n'aura le droit d'y accéder si ce collaborateur quitte définitivement l'entreprise. Ce point de vue, arc-bouté sur une disposition bien antérieure à l'usage de la messagerie électronique et incompatible avec une politique d'archivage électronique, est fort heureusement en train d'évoluer.

L'archivage électronique n'en est qu'à ses débuts et les solutions sont encore souvent partielles. Certains outils proposent un archivage limité à un type de document ou plus exactement à un type de format de fichier. Cela peut se comprendre dans la mesure où la technologie ne permet pas encore un traitement transversal de toute l'information mais il serait sans fondement de pérenniser cette approche. Archiver sur le seul critère de format informatique sans tenir compte du contenu reviendrait à agir comme ce

fonctionnaire, consciencieux mais analphabète, d'une préfecture de province dans les années cinquante que l'on avait imprudemment chargé de classer les archives et qui, après avoir tout mélangé, avait scrupuleusement rangé les registres par critère de taille et de couleur d'étiquette sans se préoccuper de la nature des informations...

L'archivage papier et l'archivage électronique ont encore ce point commun qu'ils doivent gérer conjointement la valeur de preuve et la valeur d'usage, la sécurité de l'information et sa mise à disposition, la maîtrise des contenus et la gestion des contenants.

## **Traçabilité**

---

Après son utilisation pour la gestion des données, puis pour l'accès à l'information via la numérisation, l'informatique est utilisée pour produire des actes de façon autonome sans aucune intervention du papier, ni au départ ni à l'arrivée.

L'acte électronique a été permis et même favorisé par la directive européenne du 13 décembre 1999, relayée en France par la loi du 13 mars 2000 qui pose le principe de la reconnaissance de l'écrit électronique au même titre que l'écrit papier. D'autres lois et divers décrets sont venus depuis 2000 développer ou préciser les conditions d'élaboration et de transmission du document numérique natif.

Dopées par cet arsenal législatif et réglementaire, les techniques d'archivage et de conservation cherchent et trouvent leurs marques dans l'environnement numérique, en tirant le meilleur parti des contraintes et des atouts de la technologie, tout en restant fidèles aux fondamentaux de l'archivistique.

Une question toutefois reste en suspens. Puisque la caractéristique première du numérique est de dissocier l'information de son support, comment contrôler l'authenticité et la fiabilité de l'information après sa création et son archivage ? Comment être sûr qu'elle émane bien de l'auteur annoncé ? Comment être sûr que rien ni personne n'en a altéré le message ou la date ?

Dans l'environnement papier, le document authentique est l'original, c'est-à-dire le document originel revêtu des signes observables de validation que sont une signature manuscrite, une date, un numéro d'enregistrement ou un tampon, un document dépourvu de ratures douteuses, d'incohérences de forme ou de présentation.

Jusque dans les années 1970, on trouve de beaux spécimens d'actes et de courriers ornements de multiples cachets, mentions, visas qui en disent long sur l'instruction de l'affaire. Ils en disent autant sinon plus que le texte brut qui n'inclut pas son devenir. Le texte brut est en quelque sorte mort-né. Ce sont les traces de sa vie qui permettent de reconstituer ce qui s'est réellement passé.

La bureautique et les réseaux, en dématérialisant la transmission, ont introduit le syndrome du document propre... mais muet. Or, ce que l'auditeur ou l'utilisateur a besoin de savoir, c'est qui a validé, qui a lu, qui a transmis, qui a exécuté, qui a contrôlé. Toutes ces informations se trouvent aujourd'hui dans les messages électroniques d'envoi des

documents, dans les messages instantanés de discussion sur les contenus, dans des notes prises à côté pour ne pas « dénaturer » l'original, alors que la nature même de l'original est de recueillir ces traces. Peu importe tant qu'il s'agit de documents internes ou secondaires. Mais dès lors que l'acte électronique est permis, il ne peut rester nu.

Mais comment reconnaître un original numérique puisque tous ces éléments additionnels à l'écrit, réduits eux aussi à l'état de 0 et de 1, n'offrent aucune prise à la critique formelle ?

On l'entend souvent : dans l'environnement numérique, l'original n'existe plus. A l'écran, le fichier numérique originel ne se distingue pas de la copie qui peut en être faite. La copie ne présente aucune différence sensible avec l'original comme ce peut être le cas avec une photocopie, une photographie ou une recopie manuelle. Mais de l'autre côté de l'écran ?

« L'original est mort. Vive la trace numérique » déclare Isabelle Renard[2].

Le roi est mort, vive le roi ! Le roi est toujours là ; ce n'est plus le même personnage (Louis a succédé à Henri) mais la fonction perdure. La forme traditionnelle de l'original disparaît mais le concept perdure sous une autre forme, plus contemporaine, adaptée au nouvel environnement technologique. Qu'est-ce qu'un original numérique ? Un objet constitué d'une suite de 0 et de 1, doté de caractéristiques telles qu'un juge pourra se convaincre, à tout moment, que ce qu'il dit ou semble dire (quand on le voit à l'écran) est vrai.

De même que le support de l'information prend la forme d'une succession d'éléments articulés, le caractère original de l'information numérique se manifeste au travers d'une succession de traces provoquées ou générées automatiquement par l'environnement électronique dans lequel le document ou l'objet numérique est produit, transmis, archivé, si toutefois cet environnement est conçu pour cela.

On peut distinguer, outre le message énoncé par l'auteur du document, deux ensembles d'éléments constitutifs de l'original numérique : d'abord, les données ou fichiers de signature et d'horodatage qui établissent son authenticité en attestant l'identité de l'auteur et la date du document ; ensuite, les informations enregistrées à chaque opération de nature à affecter l'information initiale : diffusion, avis de réception, tentative de modification du contenu, lecture, transfert vers un autre système, etc.

Les premières données « originalisent » le document ; les suivantes pistent cette originalité et, sauf accident, la maintiennent.

Pour être recevable à titre de preuve, l'écrit électronique doit être authentique. L'authenticité est la caractéristique initiale de l'acte, que l'environnement électronique met en péril par la dissociation des données et du support. Pour que l'authenticité existe à l'arrivée, c'est-à-dire le jour où le document est utilisé, il faut qu'elle ait été préservée, qu'elle n'ait pas été altérée, qu'elle soit intègre. Le moyen de garantir cette intégrité est la conservation et la traçabilité de tous les événements de sa vie.

## L'autopsie de l'information ou la diplomatie numérique

---

Si un doute surgit quant à l'authenticité ou la fiabilité d'une archive numérique, il faudra vérifier les traces qui lui sont attachées mais aussi, comme on le fait pour les archives traditionnelles, étudier la vraisemblance du discours, en le comparant à d'autres informations de même nature, émanant du même auteur, adressées au même destinataire ou traitant du même sujet. La falsification peut toucher aussi bien la validation de l'acte que son contenu.

Il convient donc d'expertiser la valeur de l'information. Compte tenu de l'impossibilité d'une analyse externe, directe et sensorielle de l'archive virtuelle, il s'agit plutôt d'autopsier l'objet d'information en vérifiant l'intégrité de l'ensemble puis en étudiant l'origine et l'histoire de ses composantes : leurs allées-venues, leurs connexions, leurs mutations.

Les outils d'analyse de la validité d'un document électronique ne peuvent pas être moins sophistiqués que les outils de production. Les faussaires et ceux qui les traquent ont toujours entretenu une saine émulation sur le plan des moyens. La part technologique dans l'expertise de la qualité ou de la véracité d'un original numérique est donc essentielle.

Mais la science ne suffit pas. L'exercice requiert aussi de la méthode. Comment appréhender la critique d'un document ? Comment et où a-t-il été établi ? Par qui ? Pour qui et pour quoi ? Est-il conforme aux autres documents établis par la même personne, au même endroit, à la même époque ? Son discours, son style, son allure sont-ils vraisemblables ?

A vrai dire, si les méthodes d'analyse et de critique de l'information n'avaient pas été inventées au XVII<sup>e</sup> siècle, il faudrait les inventer aujourd'hui. Comment démontrer qu'un document est faux ou sincère ? Comment décomposer, qualifier et classer les composantes de la trace écrite pour appuyer cette démonstration ? De solides réponses à ces questions ont été fournies par dom Jean Mabillon, moine bénédictin, un des plus brillants érudits de son temps, ce qui lui valut, plus d'un siècle après sa mort, en 1817, de donner son nom à une rue de Paris[3].

Tout a commencé par une querelle politico-religieuse dans laquelle un jésuite (dont l'ordre avait tout juste cent cinquante ans) mettait en cause l'authenticité de certains titres de propriété de l'abbaye Saint-Germain-des-Prés, grand, riche et séculaire propriétaire terrien. Cependant, l'accusation était insuffisamment étayée. Le plus réputé des historiens de l'abbaye, Mabillon, répondit par la méthode : il se livra à une analyse objective et comparative de l'ensemble des chartes ou *diplômes* de l'abbaye dont il étudia la forme et le contenu, l'écriture, les formules, le système de datation et les modes de validation. Les conclusions de l'historien furent suffisamment éclairantes pour mettre fin à la polémique en faveur de l'abbaye mais surtout, en 1681, Mabillon tira de ses

travaux un traité, le *De re diplomatica*, qui posa la première pierre de l'expertise en écriture et de la diplomatique, discipline fondamentale du cursus de l'École des chartes depuis sa création en 1821.

La diplomatique est définie comme la discipline qui étudie la *genèse*, la *forme* et la *tradition* des documents d'archives ; elle analyse selon quelles règles tel ou tel écrit administratif ou officiel a été établi, dans quel contexte politique et économique, dans quel but pour celui qui le signe, de quelle manière il a été validé, comment et à qui il a été transmis.

Grâce à son étude comparée des actes médiévaux, Mabillon théorisa différents aspects de la production documentaire administrative. La diplomatique examine successivement les caractères externes de l'acte (support, format, mise en page, écriture) puis ses caractères internes : l'adresse et les références, le texte (qui comporte l'exposé du contexte, la décision exprimée par un verbe d'action et les clauses), les moyens de validation, la langue et le style.

Longtemps, le champ de la diplomatique a été circonscrit à une période (le Moyen-Age) et à certains documents (les actes officiels) avant de s'étendre progressivement aux époques moderne et contemporaine puis de s'intéresser à tous les types d'archives.

Malheureusement, la diplomatique en tant que telle est restée une discipline savante, essentiellement réservée à la connaissance historique rétrospective, confidentielle en raison d'un nom sibyllin pour les non-initiés et d'un vocabulaire technique certes charmant (l'invocation, le dispositif, l'eschatocole...) mais décalé dans l'actuelle société de l'information.

On parlerait aujourd'hui de processus d'élaboration ou de production, de présentation et de format d'enregistrement, de diffusion des documents, toutes ces notions s'inscrivant dans la démarche de l'ILM (*Information Lifecycle Management*) évoquée au chapitre précédent et qui s'affirme comme la plus pertinente pour un archivage efficace des données numériques[4].

C'est réellement dommage – mais non irrémédiable – que l'apport de la diplomatique dans l'analyse de l'information ne soit pas davantage pris en compte par les professionnels de l'archivage et par les outils d'archivage électronique.

Mabillon n'est cependant pas totalement ignoré en dehors de la sphère de l'histoire médiévale. En témoignent notamment les travaux sur les archives numériques et l'archivage électronique menés par le groupe de travail international InterPARES, depuis une quinzaine d'années déjà, à l'Université de Vancouver, sous la direction de Luciana Duranti [5].

Bien sûr, l'analyse des caractéristiques externes et internes de l'information, complétée par une étude comparative, a d'abord visé la détection des faux en écriture. Cette pratique de la discipline *a posteriori* reste pertinente bien que limitée en volume d'activité

(on constate en effet que la nature du support n'influe pas véritablement sur la proportion de faussaires).

Mais ce qui est intéressant surtout, c'est l'utilisation prospective de la méthode transposée et appliquée à la production, au format et à la transmission de l'information numérique. Car c'est dans une approche prospective que la méthode diplomatique se montre aujourd'hui la plus prometteuse. Elle se révèle en effet fort efficace pour améliorer *a priori* la conception et la production de l'information qui sera archivée : il est par exemple utile de distinguer dès la création d'une application informatique quelles données devront être archivées afin d'anticiper la faisabilité de leur extraction et de leur transfert vers un système d'archivage.

C'est dans ce sens que l'on peut parler de diplomatique numérique.

L'analyse méthodique *a priori* des processus et de l'environnement de création de l'information permet de produire une information mieux structurée, mieux balisée, mieux contextualisée et, partant, plus percutante, plus sûre, plus résistante. Cette analyse se traduit en métadonnées qui correspondent à une sorte de carroyage de l'objet documentaire et à une mise en perspective de chaque parcelle d'information en fonction de ses coordonnées spatio-temporelles dans l'ensemble auquel elle appartient. Les choses étant tracées dès l'amont de la vie de l'information, celle-ci se présentera sous une forme optimale au moment de sa mise en archive. C'est le moyen d'anticiper les risques de mise en danger de l'intégrité, de l'authenticité et de la fiabilité de l'information archivée induits par la technologie numérique.

*Mutatis mutandis*, Mabillon eût-il vécu au XXI<sup>e</sup> siècle, nul doute qu'il eût reçu le César de la métadonnée et l'oscar de la traçabilité.

---

[1] Arlette Farge, *Le Goût de l'archive*, Seuil, 1989

[2] « L'original est mort, vive la trace numérique ! », par Maître Isabelle Renard, August & Debouzy, [http://solutions.journaldunet.com/0302/030227\\_juridique.shtml](http://solutions.journaldunet.com/0302/030227_juridique.shtml), 27 février 2003

[3] La personnalité de Mabillon semble toutefois étrangère au baptême de la station de métro du même nom, imputable principalement à la proximité de la rue.

[4] Jean-Marc Rietsch, Marie-Anne Chabin et Éric Caprioli, *Dématérialisation et archivage électronique. Mise en œuvre de l'ILM*, Dunod, 2006

[5] Luciana Duranti, *Diplomatics : new uses for an old science*, Scarecrow Press, Lanham and London, 1998 et [www.interpares.org](http://www.interpares.org)