

# Section 13 - Synthèse

version 1

**FRANÇOISE BANAT-BERGER**  
**CLAUDE HUC**

22 novembre 2011

# Table des matières

<b>Chapitre 1. Objet de cette section</b>	<b>3</b>
<b>Chapitre 2. Avancées constatées durant ces dernières années</b>	<b>4</b>
<b>Chapitre 3. Les obstacles restent nombreux</b>	<b>6</b>
<b>Chapitre 4. Étapes clés d'un projet d'archivage électronique en quelques lignes</b>	<b>7</b>
<b>Glossaire</b>	<b>8</b>
<b>Bibliographie</b>	<b>9</b>
<b>Webographie</b>	<b>10</b>

# Chapitre 1. Objet de cette section

Il s'agit d'une part de faire une synthèse de ce qui nous semble être des acquis en matière d'archivage numérique durant ces dernières années et des difficultés qui restent encore à franchir dans ce domaine.

Il s'agit d'autre part de résumer en quelques lignes les conditions nécessaires pour conduire et réussir un projet d'archivage numérique.

Chaque notion ou problématique évoquées renvoie bien évidemment aux différentes parties du module, auxquelles il conviendra de se reporter suivant les sujets évoqués (normalisation, formats, supports, organisation, processus, stratégies, études de cas...)

# Chapitre 2. Avancées constatées durant ces dernières années

Le chemin parcouru depuis quelques années dans le domaine de l'archivage numérique est tout à fait important, notamment dans le domaine des normes et dans la mise en œuvre de premiers projets concrets

## 2.1. Progrès de la normalisation

Les progrès en matière de normalisation ont été remarquables durant ces dernières années ; ils ont permis, en se reposant sur les normes édictées, de nets progrès dans la qualité des projets conduits, mais aussi le partage de savoirs communs entre un grand nombre d'acteurs.

A commencer par le **modèle conceptuel OAIS** :

- il a rassemblé de grands acteurs du monde scientifique et patrimonial,
- il a fourni un vocabulaire commun,
- il a donné également un modèle d'information et un modèle fonctionnel qui peut être aisément transposé dans tous les environnements et contextes pour l'ensemble des secteurs d'activité,
- il permet aussi de mettre en place une méthodologie rigoureuse indispensable à la conduite d'un projet.

Parmi les normes généralistes :

- on s'appuiera également avec profit, pour ce qui recouvre le périmètre d'une gestion raisonnée du patrimoine informationnel dans un organisme, sur la norme ISO 15489 relative au **record management** ;
- du côté des spécifications techniques concernant la mise en place d'un système d'archivage électronique (incluant les processus de numérisation des documents papier ou audiovisuels), on peut **en France se référer à la norme NF Z 42-013 dans sa dernière version de mars 2009** (à cet égard, a été accepté le principe à l'ISO d'une transposition de cette norme au niveau international).

Les progrès de la normalisation ont également touché de nombreux autres domaines intéressant l'archivage numérique et particulièrement dans le domaine des **formats** , des **métadonnées**  ou encore des **transferts**.

Dans le domaine bureautique notamment, la publication d'un format spécifique pour l'archivage, **le format PDF/A en 2005** a constitué une véritable réponse à la question délicate de l'archivage de documents bureautiques simples dont on ne maîtrise généralement pas les conditions de production et qui s'appuient le plus souvent sur des formats dont on ne connaît pas les spécifications.

De même que le mouvement de **normalisation d'un certain nombre d'autres formats** :

- comme OpenDocument Format (ODF),
- ou encore la version 1.7 de PDF,
- ou encore celle du format porté par Microsoft, Office Open XML (OOXML).

Concernant les métadonnées  on a noté l'émergence de plusieurs formats de métadonnées, avec leurs spécificités :

- par secteurs comme les métadonnées géographiques,
- par types : descriptives, d'empaquetage
- ou encore orientées pérennisation comme PREMIS.

Enfin concernant le domaine stratégique de la **prise en charge des SIP**, les spécifications et formats orientés échanges sont essentiels : la **norme PAIMAS**, le projet de **norme PAIS** ou encore **le format d'échange en France (le SEDA)**.

Enfin, les progrès en matière de mise en œuvre de plateformes d'archivage électronique sont attestés par le besoin de plus en plus pressant de déterminer des critères visant à pouvoir auditer et certifier les services et systèmes d'archivage électronique (projet de norme ISO, certification pour la gestion de l'archivage (le records management) fondée sur les spécifications, projet en France de certifications sur la base des normes liées à la sécurité des systèmes d'information, ou encore de certifications sur la base de **référentiels reposant sur la norme NF Z 42-013**).

A cela, on peut ajouter la présence dans de nombreux pays, au moins européens, de cadres communs d'interopérabilité qui soutiennent de fait les projets d'archivage électronique, en préconisant des sémantiques communes, des formats ouverts...

## 2.2. Prise de conscience qui s'affirme et projets qui se concrétisent

On observe une prise de conscience plus grande dans un certain nombre d'organismes ou de médias concernant la nécessité d'organiser une conservation pérenne des informations numériques, à la mesure de l'augmentation croissante des volumes (ainsi l'exemple du domaine aéronautique, spatial ou de la Bibliothèque nationale de France) et, dans le secteur public ou celui des entreprises, à la faveur des évolutions réglementaires qui consacrent la valeur d'originaux aux écrits numériques signés électroniquement ; de même les citoyens voient le numérique envahir l'ensemble de leurs activités et commencent à avoir des problématiques de conservation.



### Exemple

Des évolutions comme la dématérialisation des bulletins de paye désormais possible sur la base du volontariat en France, pose ce problème d'une manière aiguë : les premières réponses restent partielles et imparfaites mais on commence à voir émerger des offres de coffres-forts électroniques (par le Gouvernement avec « Mon service public » ou par des banques....).

L'archivage électronique **apparaît même modestement dans les plans** gouvernementaux, les salons sur le numérique **intègrent systématiquement la problématique** dans leur affichage même si la compréhension du problème reste imparfaite, les éditeurs développent des produits dits d'archivage électronique, de gestion des archives, de coffres-forts électroniques, même si là encore le périmètre couvert est souvent partiel. Les éditeurs documentaires, travaillant jusque là sur l'accès à des ressources sur support papier, commencent à investir le domaine du numérique. Les éditeurs de solutions d'infrastructures de stockage commencent à intégrer, pour certains d'entre eux, le modèle OAI et par conséquent à penser leurs solutions dans un environnement plus large.

On constate également une **augmentation des formations courtes** dans ce domaine, ou du moins d'une partie des sujets liés à ce domaine, que les organismes privés ou publics commencent à offrir.



### Exemple

C'est ainsi que, depuis 5 années, le groupe Pérennisation des informations numériques (PIN) a mis en place en France, une à deux fois par an, une formation de quatre à cinq jours cohérente et globale qui s'adresse à la fois à des gestionnaires de l'information et à des informaticiens.

Enfin on **peut compter** des réalisations inexistantes encore il y a **cinq années**, comme en attestent les études de cas étudiés dans la section 12.

On remarque également une **volonté de mettre en œuvre des mutualisations** : dans le domaine des archives institutionnelles entre les collectivités territoriales (mise en commun de briques logicielles, portage de projets au niveau d'une région, prestations de service d'un conseil général aux communes du département) ou encore dans le domaine des archives de la recherche et de l'enseignement supérieur avec une responsabilité d'archivage pérenne reconnue au CINES ou encore l'utilisation d'outils logiciels permettant la duplication de l'information entre plusieurs serveurs et sites répartis. Ces projets de mutualisation en restent encore à leurs débuts et cherchent à trouver de nouveaux modes d'organisation plus souples et plus réactifs que dans leur environnement habituel.

## Chapitre 3. Les obstacles restent nombreux

Si des progrès indéniables sont accomplis, il n'en reste pas moins que les difficultés restent bien réelles :

- normes parfois trop nombreuses, redondantes en partie, utilisant des vocabulaires et des concepts différents suivant la nature des organismes qui en sont à l'origine ou des secteurs d'activité, ou à l'inverse, normes qui manquent (par exemple dans le domaine des identifiants pérennes et uniques, des métadonnées scientifiques ou encore dans le domaine de registres internationaux de formats qui émergent très difficilement),
- prise de conscience des dirigeants trop faible avec une grande difficulté à concevoir la pérennisation de l'information numérique comme une activité à part entière et par conséquent à lui accorder suffisamment de moyens. Plus généralement on constate une difficulté réelle à faire prendre en considération une problématique par essence à long terme à des dirigeants de plus en plus pressés par le court terme,
- incompréhension encore très partagée des informaticiens sur ce domaine, au mieux connaissant uniquement les problématiques de stockage et de sauvegarde,
- difficile adaptation des professionnels de l'information au nouvel environnement numérique, tant dans des pratiques de records management à mettre en place que pour la claire compréhension de la problématique de la pérennisation tout en maintenant une activité de plus en plus soutenue en matière de rationalisation de l'archivage papier dans un contexte de changements des organisations de plus en plus rapides.
- difficile voire absence de dialogue entre les métiers, les informaticiens, les gestionnaires de l'information, qui conduit à des échecs dans les projets,
- manque de formations initiales et continues sur le sujet associant ces différents profils de manière à faire émerger de nouveaux métiers (de gestionnaires de l'information numérique et de gestionnaires de plateformes d'archivage numérique),
- sous-estimation généralisée des projets d'archivage numérique, l'estimation s'arrêtant aux économies induites par la diminution des locaux de stockage pour les archives papier.

## Chapitre 4. Étapes clés d'un projet d'archivage électronique en quelques lignes



### Attention

La mise en œuvre d'une plateforme d'archivage électronique n'entraîne pas forcément des investissements matériels et logiciels importants, mais elle requiert avant tout :

1. une prise de conscience du problème posé par la pérennisation des informations numériques
2. une mise à niveau de compétences variées et complémentaires
3. un projet transverse impliquant services d'archives, services informatiques, services métiers, services qualité, méthode...

On peut proposer à titre indicatif, la liste d'actions à conduire.

1. Obtenir de sa direction un soutien clair qui se manifeste par une lettre de mission, précisant qui est le chef de projet, le périmètre de son action et les modalités de collaboration des différents services.
2. Travailler en collaboration étroite avec les services informatiques de l'organisme.
3. Réunir les compétences nécessaires : archivistiques, juridiques, métiers, informatiques (structuration des données, langages XML, formats des données, stockage et sauvegarde, réseaux et sécurité informatique, applications Web) et faire en sorte que tous les acteurs partagent la même compréhension de l'archivage.
4. Etablir une politique d'archivage pour l'organisme et pour chaque grande catégorie documentaire, un protocole de versement et un contrat de service avec les producteurs et les services informatiques précisant les rôles, missions et responsabilités de chacun des partenaires ainsi que les niveaux de service attendus.
5. Avoir une stratégie en matière de gestion de l'information dans son environnement : identifier les principales sources d'information (systèmes d'information) et pour chacune d'elles, savoir suivant les durées de conservation si elles sont destinées à être gérées tout au long de leur cycle de vie au sein de l'environnement de production ou, à l'inverse, être transférées, à un moment à déterminer, catégorie par catégorie, vers une plateforme d'archivage numérique et avoir ainsi quantifié et qualifié les archives à transférer.
6. Définir les modalités de destruction des archives si, au terme de leur cycle de vie, celles-ci sont destinées à être détruites et la traçabilité de ces destructions.
7. Porter une attention particulière aux données à caractère personnel qui requièrent généralement un traitement spécifique.
8. Avoir une vision globale et unifiée portant sur le papier et le numérique afin d'assurer un traitement homogène au niveau de l'accès.
9. Avoir une connaissance précise des utilisateurs et de leurs besoins en termes de critères de recherche et d'accès.
10. Etablir ou utiliser un format donné pour le transfert de données (SIP) et étudier le format des AIP et des DIP.
11. Réfléchir aux problématiques de nommage et d'identifiants.
12. Définir sa stratégie en matière de formats acceptés par le système et de conversion si nécessaire. Etudier les réalisations et expériences dans le domaine et les partager autant que possible. Déterminer les modalités de récupération et d'exploitation des métadonnées techniques relatives aux formats de fichiers reçus.
13. Définir sa stratégie en matière d'archivage de fichiers signés : outils de vérification des signatures, de production de rapports de vérification et d'archivage de ces rapports.

# Glossaire

## Format

Un format, dans son sens le plus général, permet de définir les caractéristiques physiques ou logiques d'un support d'information. Les formats sont le plus souvent normalisés ou standardisés.

Le format peut définir :

- soit le support physique, on parlera alors de format de support ; dans ce cas, il précisera les caractéristiques physiques de ce support : A4 est un format papier de dimensions 21cm x 29.7 cm,
- soit les caractéristiques logiques d'organisation de l'information, nous parlerons alors de format de données ou de format de fichiers (voir la définition),
- soit l'ensemble des caractéristiques physiques et logiques qui peuvent être imbriquées (VHS, CD-Photo Kodak), situation peu propice à la pérennisation.

## Métadonnées

Étymologiquement, « méta » provient du grec signifiant « après, au-delà de, avec » : « méta » données signifie « au-delà des données », « qui dépasse les données », « qui englobe les données ». Les métadonnées sont donc des données sur les données, à propos des données, qui définissent, décrivent des données, leur contexte, leur contenu, leur structure des ainsi que leur gestion dans le temps.

# Bibliographie

[Premier ouvrage de synthèse sur l'archivage numérique en langue française.] • BANAT-BERGER F., HUC C., DUPLOUY L., L'Archivage numérique à long terme, les débuts de la maturité? Paris, La Documentation française, 2009.

# Webographie

[Norme de référence essentielle pour comprendre le problème posé par l'archivage numérique]  
[http://public.ccsds.org/publications/archive/650x0b1\(F\).pdf](http://public.ccsds.org/publications/archive/650x0b1(F).pdf)