

---

**PARKER WILLIBORG**

---

## **Direction des Archives de France**

**Etude relative à la réalisation de plates-  
formes d'archivage électronique pour les  
services publics d'archives**

**Etude de l'existant et des besoins**

Version 1.1 – 14 Mars 2005

# Sommaire

<b>1. CONTEXTE ET OBJECTIFS</b> .....	<b>3</b>
<b>2. DÉROULEMENT DE LA PHASE D'ENTRETIENS</b> .....	<b>4</b>
<b>3. FAITS MARQUANTS ASSOCIÉS À L'ARCHIVAGE PHYSIQUE</b> .....	<b>5</b>
3.1 ARCHIVES INTERMÉDIAIRES ET ARCHIVES DÉFINITIVES .....	5
3.2 VERSEMENTS D'ARCHIVES .....	5
3.2.1 Services versants .....	5
3.2.2 Nature des archives versées .....	9
3.2.3 Processus de versement aux archives .....	9
3.3 RESTITUTIONS D'ARCHIVES CONTEMPORAINES .....	11
<b>4. FAITS MARQUANTS ASSOCIÉS AUX PLATES-FORMES D'ARCHIVAGE ÉLECTRONIQUE OPÉRATIONNELLES</b> .....	<b>13</b>
4.1 NATURE ET VOLUME DES DONNÉES VERSÉES .....	13
4.2 ORGANISATION .....	14
4.3 PROCESSUS DE VERSEMENT ET DE CONTRÔLE .....	15
4.4 PROCESSUS DE RESTITUTION .....	15
4.5 INFRASTRUCTURE LOGISTIQUE DU SITE D'HÉBERGEMENT DE LA PLATE-FORME .....	16
4.6 FORMATS DE CONSERVATION .....	16
4.7 SUPPORTS DE STOCKAGE .....	16
4.8 PLATE-FORME CDC ZANTAZ (TIERS ARCHIVEUR) .....	17
<b>5. FAITS MARQUANTS ASSOCIÉS AUX PROJETS D'ARCHIVAGE ÉLECTRONIQUE IDENTIFIÉS</b> .....	<b>18</b>
5.1 PROJETS IDENTIFIÉS .....	18
5.2 CATÉGORISATION DES PROJETS IDENTIFIÉS .....	19
5.2.1 L'archivage de documents numérisés par le service producteur .....	19
5.2.2 L'archivage de documents électroniques natifs détenus par les services producteurs .....	21
5.2.3 L'archivage de données contenues dans des applications .....	22
5.3 SOURCES D'INFORMATIONS CANDIDATES À L'ARCHIVAGE ÉLECTRONIQUE .....	25
5.4 PROJETS TYPES .....	26
5.4.1 Bureau des nationalités de la préfecture du Nord : numérisation des dossiers de demandes de carte de séjour (AD59) .....	26
5.4.2 Conseil général des Yvelines : Délibérations et contrôle de légalité (AD78) .....	27
5.4.3 Juridictions pénales : la nouvelle chaîne pénale (NCP) .....	28
5.4.4 Direction de l'urbanisme de la mairie de Paris : données relatives aux permis de construire (AD75) .....	29

# 1. Contexte et objectifs

---

Ce document s'inscrit dans le cadre de la première phase de l'étude menée par Parker Williborg pour définir les fonctionnalités et évaluer les coûts associés à la mise en place d'une plate-forme d'archivage électronique générique.

Cette plate-forme s'adresse aux services d'archives publics (nationaux et locaux) ainsi qu'aux administrations productrices désireuses de disposer d'une telle plate-forme pour leurs propres besoins.

Cette première phase consiste à prendre connaissance :

- des moyens et modes d'organisation actuels des services d'archives publics, que ceux-ci aient ou non déjà mis en place une plate-forme d'archivage électronique
- du fonctionnement des plates-formes d'archivage électronique actuellement en service
- des projets d'archivage électronique déjà menés ou en cours de réflexion.

A cette fin, Parker Williborg a organisé des entretiens avec différents services d'archives. Ces entretiens ont permis d'aborder les sujets suivants :

- Pour les services d'archives non équipés à ce jour d'une plate-forme d'archivage électronique, les questions posées ont porté sur les thèmes suivants :
  - o Organisation et effectifs des centres
  - o Volumes d'archives stockés
  - o Parties versantes
  - o Flux entrants (volumes / fréquences)
  - o Flux sortants (volumes / fréquences)
  - o Natures d'archives gérées (intermédiaires, définitives), proportions
  - o Description des procédures et des instruments de recherche utilisés
  - o Présentation des logiciels de gestion d'archives utilisés
  - o Fonctionnement des moyens de restitution : salle de lecture, ...
  - o Description des infrastructures de communication et de stockage existantes dans le service d'archives et/ou au sein de son administration de tutelle
  - o Archivage de données électroniques. Projets identifiés, provenance des demandes, nature (Bases de données, documents), volumes envisagés
  - o Si le service d'archives numérise ou fait numériser des documents qu'il conserve, de quels équipements dispose-t-il ? Comment conserve-t-il les fichiers ainsi produits ?
- Pour les centres d'archives déjà équipés d'une plate-forme d'archivage électronique, les thèmes complémentaires suivants ont également été abordés :
  - o Historique et enjeux du projet d'archivage électronique
  - o Organisation du service d'archivage électronique. Compétences requises.
  - o Nature des informations gérées. Volumes associés.
  - o Format des données
  - o Gestion des méta-données associées
  - o Modes de versement et de restitution
  - o Modes de conservation et de contrôle
  - o Moyens informatiques mis en œuvre (logiciels, matériels, réseau)

Ces entretiens ont également permis de collecter différents documents relatifs à l'activité des services et à leur organisation.

Le présent document constitue le rapport d'étude de l'existant et des besoins issu des informations collectées lors de ces différents entretiens.

## 2. Déroulement de la phase d'entretiens

Les entretiens avec les services d'archives ont été menés durant la période du 31 Janvier 2005 au 23 Février 2005.

Les services d'archives rencontrés ont été les suivants :

Etablissement/ Service/ Organisme	Personnes rencontrées	Date de l'entretien
<b>Services d'archives nationaux</b>		
Centre des Archives Contemporaines	M. Teil Mme Petillat	11-févr
<b>Services d'archives départementaux</b>		
AD Ille et Vilaine	M. Le Clerc Mme Mikes	31-janv
AD Haute-Garonne	Mlle Regagnon	01-févr
AD Yvelines	Mme Laude M. Guérin M. Lambert (CG) M. Doublier (CG)	08-févr
AD Paris	M. Faivre d'Arcier	18-févr
AD Loire	Mme Bidou M. Carron	22-févr
<b>Services d'archives communaux</b>		
AM Lyon	Mme Marin Mme Dormont	22-févr

En complément à ces entretiens avec des services d'archives publics, d'autres entretiens ont été organisés avec d'autres établissements du fait :

- soit de l'existence d'une plate-forme d'archivage électronique opérationnelle (CNES, CDC Zantaz),
- soit d'incidences potentielles importantes avec d'autres projets (CDC FAST).

<b>Autres établissements</b>		
CNES	M. Huc	01-févr
CDC Zantaz	M. du Boullay	04-févr
CDC FAST	M. Bordier	18-févr

D'autres entretiens seront organisés durant la suite de l'étude, notamment avec certains services producteurs, afin de disposer d'informations complémentaires sur les projets d'archivage électronique identifiés considérés comme les plus pertinents pour notre étude.

## 3. Faits marquants associés à l'archivage physique

---

### 3.1 Archives intermédiaires et archives définitives

Les différents services d'archives publics rencontrés, qu'ils soient municipaux, départementaux ou nationaux, ne gèrent dans leur établissement que des versements d'archives définitives.

Quelques établissements conservent encore des archives intermédiaires (AD42 et AD31) mais n'acceptent plus de nouveaux versements.

Ce constat s'explique d'une part par le niveau de saturation atteint par un grand nombre de centres rencontrés et d'autre part par la circulaire du premier ministre du 2 novembre 2001 motivant la conservation et l'organisation des archives intermédiaires dans les locaux des services producteurs.

Les services d'archives publics interviennent néanmoins auprès des services producteurs pour les conseiller dans la gestion de ces archives intermédiaires, établir les tableaux de gestion, contrôler les conditions de conservation, valider les demandes de destructions ...

Les archives intermédiaires sont dans la plupart des cas éclatées dans les différentes entités du service producteur.

Dans certains cas, un centre de pré-archivage a été mis en place dans un ou plusieurs services producteurs sous la responsabilité du service avec l'assistance du service des archives. Ces centres restent peu nombreux et relativement récents. Ils font suite soit à une situation d'urgence (découverte d'archives en zone inondable à Paris) soit à une réorganisation physique d'un service producteur (Municipalité de Lyon). Seul le centre de pré-archivage de Paris est géré directement par les archives départementales de Paris.

### 3.2 Versements d'archives

#### 3.2.1 Services versants

Il s'avère difficile de se faire une idée concrète de l'importance relative des services producteurs dans les versements effectués chaque année dans les services d'archives. En effet :

- Nous n'avons pas toujours pu collecter d'informations sur la décomposition des fonds W par service producteur
- La majeure partie des centres que nous avons visité étaient saturés et ont par conséquent limité autant que possible les versements ces dernières années
- Les services producteurs ne versent pas périodiquement leurs archives. Les versements sont souvent liés aux capacités de stockage du service producteur. Il est par conséquent difficile de dégager une tendance à moins d'analyser leurs versements sur une période de 5 à 10 ans.

Néanmoins, les entretiens ainsi que l'analyse des rapports d'activité collectés (AD35, AD42, AD78 et AD 31) permettent de dresser, pour chaque type de centre, la liste des principaux services producteurs :

• **Services producteurs des services d'archives départementaux**

Les services d'archives départementaux gèrent les versements d'archives :

- Des collectivités locales implantées dans le départementaux à l'exception de celles disposant de leur propre service d'archives (mairies de plus de 2000 habitants, conseils régionaux ...) : Il s'agit notamment des conseils généraux et des petites communes,
- Des services déconcentrés de l'Etat implantés dans le département : préfecture et sous-préfecture, justice, inspection académique, Affaires sociales, Equipement ...
- Des établissements publics implantés dans le département
- Des offices notariaux

On dénombre en moyenne 200 services producteurs par service d'archives départementales.

L'analyse des versements récents permet de définir l'importance de ces différentes catégories d'administrations :

	AD35 2003 (ml)	AD78 2003 (ml)	AD42 2003 (ml)	AD31 2003 (ml)	Totaux	%
Services du conseil général et établissements publics départementaux	11,70	4,40	48,65	471,6	536,35	24,1%
Services déconcentrés et établissements publics de l'Etat	225,37	540,00	621,43	298,54	1685,34	75,9 %
<b>Total (en ml)</b>	<b>237,07</b>	<b>544,40</b>	<b>670,08</b>	<b>770,14</b>	<b>2221,69</b>	<b>100%</b>
<b>Nb de versements</b>	<b>32</b>	<b>43</b>	<b>20</b>	<b>41</b>	<b>107</b>	
<b>Moyenne des versements (en ml)</b>	<b>7.41</b>	<b>12.66</b>	<b>33.50</b>	<b>18.78</b>	<b>20.76</b>	
<b>Population (en nb d'habitants)</b>	<b>903.400</b>	<b>1.354.000</b>	<b>728.524</b>	<b>1.046.000</b>	<b>4.555.400</b>	
<b>Versement pour 1000 habitants (en ml)</b>	<b>0,26</b>	<b>0,40</b>	<b>0,92</b>	<b>0,74</b>	<b>0,49</b>	

On constate que la part des versements effectués par les services déconcentrés de l'Etat apparaît comme largement majoritaire : 75.9 %.

Parmi ces services déconcentrés, les principaux services versants sont :

- Ceux dépendant du ministère de la Justice (juridictions)
- Ceux dépendant du ministère de l'Intérieur (préfectures, sous-préfectures, commissariats)

**Les juridictions** procèdent à des versements relativement réguliers et leurs versements sont volumineux et très homogènes puisqu'il s'agit pour l'essentiel des minutes des jugements ainsi que des dossiers de procédures.

Classement des plus gros services versants dépendant du ministère de la Justice :

	AD31 2004	AD35 2003	AD78 2003	AD42 2003	AD42 2002

N° 1	Cour d'appel de Toulouse	Chambre régionale des comptes	Cour d'appel de Versailles	TGI de Saint Etienne	TGI de Saint Etienne
N° 2	Dir. Dép. de la protection judiciaire	Tribunal administratif de Rennes	TGI de Versailles	Conseil des Prud'hommes de Roanne	TGI de Montbrison
N° 3	Centre de détention de Muret	Cour d'appel de Rennes	Tribunal de commerce		
N° 4	Dir. Rég. De la protection judiciaire	Conseil des prud'hommes de Saint-Malo	TI de Poissy		
N° 5	TGI de Toulouse	Conseil des prud'hommes de Rennes	Tribunal administratif		

**Les versements des services de la préfecture** sont réalisés par service. En moyenne, une dizaine de services distincts de la préfecture procèdent à des versements réguliers.

Contrairement aux juridictions, les versements effectués sont de nature différente d'un service à un autre. Les services des élections par exemple versent principalement les professions de foi des candidats, des procès-verbaux, des listes d'émargement ; les services de la circulation versent quant à eux les dossiers relatifs aux retraits de permis ; les services de l'urbanisme versent des dossiers relatifs aux permis de construire ...

Classement des plus gros services versants dépendant du ministère de l'Intérieur :

	<b>AD31 2004</b>	<b>AD35 2003</b>	<b>AD78 2003</b>	<b>AD42 2003</b>	<b>AD42 2002</b>
N° 1	Cabinet du préfet	Cabinet du préfet	Sous-préfecture de Saint-Germain en Laye	Sous-préfecture de Roanne	Direction du contrôle de légalité
N° 2	Cabinet du commissaire de la république	Direction de la réglementation	Commissariat de Houilles	Commissariat central de Saint Etienne	Sous-préfecture de Montbrison
N° 3		Direction des relations avec les collectivités locales	Direction des actions de l'Etat	Direction des affaires économiques	Direction des affaires économiques
N° 4		Direction des actions de l'état		Direction du contrôle de légalité	Direction de la réglementation
N° 5				Direction de la réglementation	Commissariat central de Saint Etienne

**Parmi les services du conseil général**, les principaux services versants sont les services sociaux. Les dossiers versés par ces services correspondent à des dossiers d'aides classés par type d'aide et / ou foyer / personne concernée.

Parmi les autres principaux services versants des conseils généraux, on compte notamment les services de l'éducation, des finances et des marchés publics.

Tout comme pour les préfetures, la nature des dossiers dépend du service versant.

Classement des plus gros services versants des conseils généraux et établissements publics départementaux :

	<b>AD31 2004</b>	<b>AD35 2003</b>	<b>AD78 2003</b>	<b>AD42 2003</b>	<b>AD42 2002</b>
N° 1	Direction des finances	Pôle éducation, jeunesse, culture et sport	Direction de l'action sociale	Direction de l'éducation	Direction de l'éducation
N° 2	Direction de la solidarité			Direction de la protection sociale	Direction de la voirie
N° 3	Direction des marchés publics			Direction aménagement rural	Direction aménagement rural
N° 4	Direction des interventions			Direction de la voirie	Direction des transports
N° 5	Service du personnel			Direction des affaires économiques	Direction des affaires économiques

- **Services producteurs des services d'archives municipaux**

Les services d'archives municipaux gèrent les archives de leur municipalité et des établissements publics municipaux.

Les principaux services versants identifiés à Paris et Lyon sont les services de l'urbanisme, du personnel et des actions sociales.

Les services de l'urbanisme versent notamment les dossiers de permis de construire clos qui sont directement communicables au public et fortement consultés en salle de lecture.

Les autres principaux services versants sont les services de l'état-civil, du patrimoine (dossiers de propriété), des marchés, des finances (budgets, politique financière, gestion de la dette), funéraires (gestion des concessions et cimetières) et de l'éducation (cantines, assistantes maternelles), ainsi que les cabinets du maire et des élus (délibérations, arrêtés).

Les principaux établissements publics municipaux versants sont les établissements culturels ainsi que les centres communaux d'assistance sociale (CCAS).

- **Services producteurs du centre des archives contemporaines**



Le centre des archives contemporaines (CAC) gère les archives des ministères et des établissements publics à compétence nationale à l'exception des ministères de la Défense, des Affaires étrangères et des Finances.

Le CAC s'appuie pour cela sur les missions des archives nationales implantées dans les ministères.

En 2003, 4,4 km d'archives ont ainsi été versées, correspondant à 598 versements et 19317 articles.

Elles se répartissent de la façon suivante :

- 376 ml d'archives de cabinets ministériels
- 3249 ml d'archives administratives
- 633 ml d'archives d'établissements publics nationaux
- 211 ml d'archives privées

Les principaux services versants sont les ministères de la justice et de l'intérieur.

### 3.2.2 Nature des archives versées

L'essentiel des archives versées se composent de dossiers composites.

Il s'agit soit de dossiers sériels (produits dans le cadre d'une procédure spécifique) portant sur des personnes ou des affaires ; soit de dossiers de principe (élaboration d'une réglementation, dossiers thématiques...) Il s'agit de dossiers composites dans la mesure où les documents contenus dans ces dossiers peuvent être de nature très différentes. Il pourra s'agir :

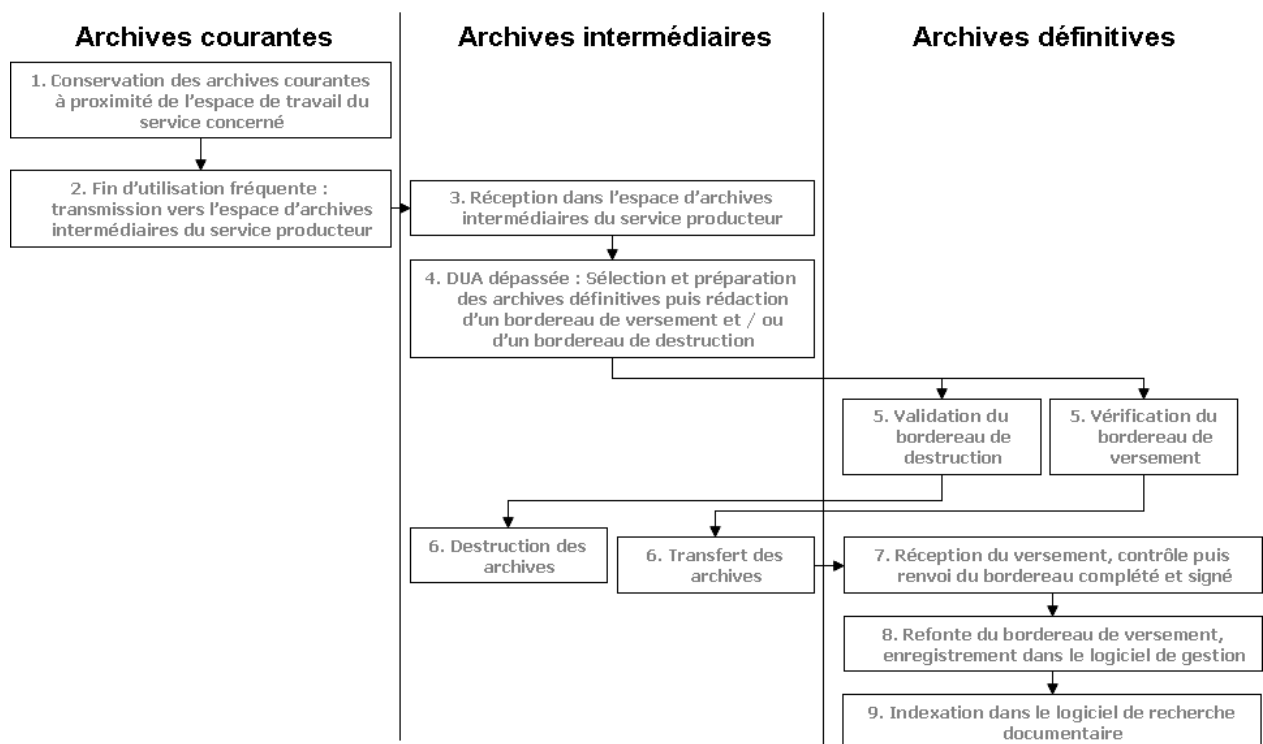
- d'actes signés,
- de pièces justificatives,
- de correspondances entrante et sortante de l'administration concernée,
- d'état imprimés issus d'une application.

Les dossiers comportant un contenu homogène sont relativement rares.

On distingue d'autre part les registres qui, en dehors de leur propre intérêt historique, sont également utilisés par les services d'archives pour retrouver plus facilement un dossier ou un document dans un lot.

### 3.2.3 Processus de versement aux archives

Le processus de versement aux archives est quasiment identique pour l'ensemble des services rencontrés. Celui-ci peut être synthétisé par le diagramme suivant :



L'espace d'archives intermédiaires peut être spécifique au service versant ou commun avec d'autres services, généralement rattachés à la même administration (mairie, préfecture, conseil général).

## 3.3 Restitutions d'archives contemporaines

### • Profil des chercheurs

On distingue trois catégories de chercheurs :

- les services producteurs : leurs recherches portent exclusivement sur leurs propres fonds.
- les chercheurs « professionnels » (étudiants, historiens, scientifiques) : leurs recherches portent sur des sujets précis et donc sur des fonds spécifiques rarement consultés. Ils peuvent dans certains cas disposer de dérogations leur permettant d'accéder à des fonds non communicables.
- les chercheurs « amateurs » type généalogistes : leurs recherches portent généralement toujours sur les mêmes fonds.

### • Nature et volume des restitutions

Peu de statistiques existent sur les consultations spécifiques aux archives contemporaines.

L'essentiel des consultations porte sur les fonds anciens et modernes. Les consultations des archives contemporaines restent aujourd'hui très réduites, du fait notamment du délai de communicabilité qui leur est appliqué.

En 2003, aux archives d'Ile et Vilaine, on comptait 2.409 articles restitués issus de la série W sur 31.809 au total soit 7,6%.

Les restitutions aux services producteurs sont les plus fréquentes. Le service producteur s'appuie pour cela sur les exemplaires des bordereaux de versement qu'il a en sa possession. La consultation des documents se fait alors soit en salle de lecture soit par l'envoi du ou des documents recherchés au service.

Les consultations réalisées par le public en salle de lecture restent assez rares excepté pour quelques fonds spécifiques (permis de construire, jugements).

Certains fonds se composent de documents dont les délais de communicabilité sont différents. Le service d'archives doit avoir veillé à ce que les documents demandés sont bien communicables avant de les restituer.

En 2003, le Centre des Archives Contemporaines :

- A reçu en salle de lecture 568 lecteurs ayant consulté 8276 articles (7477 communications). Les communications portent dans l'essentiel des cas sur 1 seul article. 227 chercheurs disposaient d'une dérogation. 15 fonds ont été consultés au moins 100 fois.
- A procédé à 9.399 communications auprès des services versants. Les principaux services concernés étant le ministère de l'intérieur (1999 communications), le service des naturalisations (1927 communications) et le ministère de la Justice (1521 communications).

Concernant les autres services d'archives, aucune information n'est disponible concernant les communications réalisées sur les archives contemporaines.

Dans les services d'archives municipales, le fonds le plus souvent consulté est celui des permis de construire.

Le nombre de chercheurs disposant d'une dérogation et le nombre d'articles communiqués par dérogation sont variables en fonction du type de service :

- 227 sur 568 pour le CAC en 2003 (40%)
- 60 sur 2424 pour les archives de la Loire en 2003 (2.5%) : 609 articles restitués
- 78 sur 3593 pour les archives d'Ille et Vilaine en 2003 (2.2 %) : 8007 articles communiqués

- **Instruments de recherche**

La recherche des documents s'effectue en premier lieu à l'aide des informations comprises sur les bordereaux de versement.

Ces informations sont généralement enregistrées dans le logiciel de gestion des versements du service puis mises à disposition pour la recherche. Suivant les cas, ces informations ne sont consultables que par les agents du service d'archives ou également par le public en salle de lecture. Les services producteurs conservent également un exemplaire de ces bordereaux qu'ils utilisent lorsqu'ils souhaitent se voir restituer un ou plusieurs articles.

Les instruments de recherche peuvent être de plusieurs natures : bordereaux de versement reçus des services producteurs et simplement complétés par les services d'archives puis conservés tels quels (collection des bordereaux, généralement classés par services producteurs) ; refonte manuelle des bordereaux pour établissement de répertoires numériques détaillés ; bases de données créées par les services d'archives à partir des informations de ces bordereaux, complétées par les services d'archives (indexation). Par ailleurs, les registres papier collectés par les services d'archives peuvent également servir pour retrouver une affaire particulière, repérer une information à un niveau de précision supérieur à celui qu'on trouve dans le bordereau de versement.

## 4. Faits marquants associés aux plates-formes d'archivage électronique opérationnelles

---

Trois centres hébergeant une plate-forme d'archivage électronique ont été visités :

- Le CNES
- Le Centre des Archives Contemporaines
- Le site informatique de la Caisse des Dépôts et Consignations d'Arcueil

Ces trois centres ont mis en place une infrastructure et une organisation permettant la conservation à long terme de données électroniques en grandes quantités et sont opérationnels depuis plusieurs années.

### 4.1 Nature et volume des données versées

Les différents centres visités sont à même d'assurer la conservation de toutes formes d'informations électroniques : données et documents.

#### - L'archivage de données

L'archivage de données représente l'essentiel des informations conservées par le CNES et par CONSTANCE.

- Le CNES

Le CNES archive les données scientifiques produites par le centre dans le cadre de projets ou d'études spatiales. Il peut s'agir de résultats de mesures ou de données résultats de calculs scientifiques.

L'archivage s'effectue essentiellement en fin de projet. Les données sont conservées d'une part parce qu'elles peuvent à nouveau servir à l'équipe projet émettrice et d'autre part car elles peuvent être réutilisées par d'autres équipes projets issues du centre ou d'autres communautés scientifiques et techniques traitant des mêmes sujets.

La qualification et la mise en forme des données à archiver sont à la charge de l'équipe projet. Des normes communes à l'ensemble des services sont en cours d'élaboration.

Ces données sont conservées entre 5 et 25 ans.

EN 2004, la plate-forme d'archivage recensait 2,1 millions de fichiers représentant un espace de stockage de 47,5 To de données (soit 22,6 Mo de moyenne par fichier).

Si on excepte les gros fichiers conservés sur un serveur spécifique, on recense 1,8 millions de fichiers représentant 7,5 To de données (Soit 4,2 Mo de moyenne par fichier).

- Constance

Constance archive les données produites par certaines administrations nationales et en particulier par l'INSEE.

Ces données sont versées sous forme de fichiers texte. Un versement peut se composer de plusieurs fichiers texte correspondant à des tables liées les uns aux autres et issues d'une même application. Une documentation technique précise doit donc également être fournie décrivant le contenu de chaque fichier et les relations entre ces fichiers.

Fin 2004, la plate-forme d'archivage recensait 5328 fichiers représentant 300 Go de données (soit 56,3 Mo de moyenne par fichier).

#### - **L'archivage de documents**

Le CNES et CONSTANCE archivent également des documents sous forme de fichiers images :

Il s'agit :

- pour le CNES des images produites par le satellite SPOT (2,4 Millions de fichiers représentant 300 To soit 125 Mo par fichier),
- pour l'INSEE des images des formulaires établis lors du dernier recensement de la population (220 Millions de fichiers représentant 14 To soit 0,06 Mo par fichier).

On constate que l'archivage de documents est considérablement plus coûteux en espace de stockage que l'archivage des données puisque dans ces deux cas, la prise en charge d'une seule source a nécessité des capacités de stockage beaucoup plus importantes que pour toutes les autres.

## **4.2 Organisation**

Au CNES, on distingue trois types d'acteurs :

- L'équipe informatique composée de 3 personnes à temps plein pour la plate-forme de stockage (STAF) : son rôle consiste à assurer la préservation des données archivées en vue de leur restitution. Cette équipe s'assure principalement de la qualité du support (contrôle, migration) et procède ponctuellement à des modifications de formats de données mais n'intervient en aucun cas sur le contenu versé.
- L'équipe valorisation des données qui définit les règles de forme et de format à appliquer lors des versements et qui contrôle certaines données versées.
- L'équipe en charge de développer les interfaces de recherche et de consultation des données archivées.
- Les responsables de projets : ils réalisent les versements de données.

En ce qui concerne CONSTANCE, on retrouve le même découpage :

- Une compétence informatique en charge de la préservation de l'intégrité des données ainsi que des chantiers de migration de support et de conversion de formats
- Une compétence archives chargée de contrôler la conformité des données versées avec les documentations fournies
- Une compétence mixte « archiviste / informaticiens » (Jean-Pierre Teil) chargée de conseiller les services producteurs dans leur démarche d'archivage électronique.

## 4.3 Processus de versement et de contrôle

Dans le cas du CNES, les versements s'effectuent via le réseau informatique haut débit du centre, la disponibilité de ce réseau haut débit ainsi que la taille importante de certains fichiers rendant ainsi l'opération plus simple que par CD-ROM.

Le versement n'est pas systématiquement contrôlé : il est de la responsabilité de l'équipe projet émettrice de faire le nécessaire en amont pour que ses fichiers restent exploitables durant la période de conservation prévue.

Dans le cas de CONSTANCE, le versement s'effectue la plupart du temps par l'envoi d'un CD-Rom comprenant les données à archiver et accompagné d'une documentation précise généralement sur support papier sur leur organisation et leur signification.

Durant l'année 2003, 5 versements de données ont été effectués :

- Présidence de la république : fichier d'enregistrement du courrier du premier septennat Chirac (29 fichiers représentant quelques Mo de données)
- Présidence de la république : fichier des invitations aux cérémonies (1 fichier de type bureautique représentant quelques dizaines de Mo de données)
- Ministère de l'agriculture : recensement général de l'agriculture (4 fichiers représentant environ 0,5 Go de données)
- INSEE : recensement général de la population (environ 7 fichiers par département soit 700 fichiers au total représentant 40 Go de données)

Fin, 2004, une vingtaine d'opérations d'archivage étaient en cours, pour le compte d'une dizaine d'organismes différents (ministère de l'agriculture, Elysée, Ined, Inserm, Ministère de l'intérieur ...).

Avant d'être versés sur la plate-forme, l'équipe Constance procède à un contrôle systématique et complet des données récupérées. Il vérifie notamment que la structure des données fournies est cohérente avec la documentation. Les principales tâches associées au traitement d'un versement sont les suivantes :

- recopie physique des fichiers
- reconstitution de leur structure pour vérification de sa validité
- tests de cohérence de la documentation avec les enregistrements fournis
- élaboration des instruments de recherche spécifiques : fiche d'application, fiche de structure, dictionnaire des données, dictionnaire des codes.

Les versements effectués sont ensuite enregistrés dans le logiciel de gestion des versements du centre, à savoir Priam.

## 4.4 Processus de restitution

Dans le cas du CNES, la restitution des données peut se faire, pour une partie importante des données archivées, directement en ligne à partir du réseau informatique du centre ou depuis Internet après s'être authentifié. Cette restitution en ligne se justifie dans la mesure où :

- les délais de restitution doivent être courts

- les utilisateurs potentiels peuvent être distants
- les volumes restitués sont très importants (10,5 To de données restituées en 2004)

Dans le cas de CONSTANCE, les restitutions sont quasiment inexistantes, à l'exception de demandes très ponctuelles faites par le service versant. Cela s'explique dans la mesure où une grande partie des informations versées ne sont pas communicables pour l'instant.

## 4.5 Infrastructure logistique du site d'hébergement de la plate-forme

Les centres visités ont mis en place, sur le site hébergeant la plate-forme, des moyens logistiques spécifiques visant à protéger les données contre des risques extérieurs et notamment les intrusions physiques, les incendies, les inondations et les coupures d'électricité.

Ces mesures de protection se matérialisent notamment par :

- le choix de locaux non inondables
- un système de sécurité comprenant badges d'accès et vidéo-surveillance
- un équipement de type « salle blanche » comprenant notamment des équipements ignifugés, des équipements anti-incendie, des onduleurs ...

## 4.6 Formats de conservation

Pour Constance comme pour le CNES, les données conservées sont essentiellement stockées sous format texte ou image.

Constance a dû procéder, dans le cadre d'un projet de migration de support, à la conversion de format des données du format EBCDIC au format ASCII.

Cette opération a été effectuée par lots de fichiers en fonction de la qualité de la documentation fournie (documentation complète, documentation partielle mais suffisante, documentation insuffisante). Les fichiers dont la documentation s'est avérée insuffisante n'ont pas pu être migrés.

Un contrôle des nouveaux supports effectué plusieurs mois après la migration a mis en évidence des anomalies de migration liées à des défauts des outils de migration utilisés. Plusieurs fichiers se sont ainsi révélés inutilisables sur le nouveau support sans qu'il soit possible de procéder à une nouvelle migration à partir des supports initiaux.

D'une manière générale, les migrations de format sont ponctuelles et font l'objet de projets à part entière. Constance s'appuie pour cela sur la notion de DUT (Durée d'Utilisation technologique) qui associe à chaque type de format une date à laquelle les fichiers conservés au format en question devront faire l'objet d'une étude de pérennité suivie éventuellement d'une opération de migration.

## 4.7 Supports de stockage

CONSTANCE a depuis le lancement du projet utilisé différentes générations de supports.



Les premiers supports étaient des bandes magnétiques. Elles ont été ensuite remplacées par des disques optiques puis des cartouches DLT (bande magnétique) et enfin des cartouches SDLT (même technologie mais capacité plus importante).

Les supports SDLT actuellement utilisés ont une capacité de stockage unitaire de 150 Go (contre 70 Go pour les DLT).

Avant la prise en charge des images issues du recensement 1999, environ 25 cartouches DLT étaient suffisantes pour stocker l'ensemble des données archivées. La prise en charge de ces images porte le nombre de supports conservés à environ 300 cartouches SDLT.

Au CNES, le choix du support dépend de la taille des fichiers versés et des attentes des services versants en terme de qualité de service et de délai de restitution. Il s'agit suivant les cas de cartouches ou de disques magnétiques.

## 4.8 Plate-forme CDC Zantaz (tiers archiveur)

Plusieurs sociétés ont développé une activité de tiers archiveur permettant l'externalisation des archives électroniques. C'est notamment le cas de la société CDC Zantaz qui a mis en place une plate-forme spécifique destinée à recevoir, stocker et restituer les données versées par leurs clients.

La plate-forme de CDC Zantaz a été en premier lieu conçue pour assurer l'archivage des mails mais elle est également utilisée pour archiver tout autre type de documents électronique.

CDC Zantaz s'appuie sur les infrastructures logistiques et informatiques de la Caisse des Dépôts et Consignations à Arcueil. Ces moyens permettent d'offrir un niveau de sécurité exceptionnel puisque ce site, qui gère déjà les traitements informatiques liés à l'activité bancaire du groupe, est protégé contre des risques de tout ordre (intrusions, inondations, incendies, coupures électriques ...) à l'aide d'équipements spécifiques : caméras et badges, reconnaissance biométrique, détecteurs de fumée, groupes électrogènes de secours ...

Le versement des données s'effectue en ligne, la plupart des cas à partir d'une connexion VPN. Le système contrôle la réception complète des données et procède à l'horodatage du versement et au calcul de l'empreinte du lot afin de garantir par la suite son intégrité.

L'archivage s'effectue sur des baies de disques magnétiques. Les données archivées sont répliquées sur 3 sites distants (dont un de back-up réalisé sur bandes). La capacité actuelle de la plate-forme est d'environ 1 To, capacité bien supérieure aux versements réalisés à ce jour. Cette capacité est largement évolutive et pourra donc évoluer dans le temps sans modification de l'infrastructure.

La restitution des données s'effectue en ligne sans délais de réponse. Une fois une requête reçue, le système recherche les documents correspondants et contrôle l'intégrité du document à l'aide de l'empreinte.

CDC Zantaz ne réalise aucune intervention sur le contenu même des versements et ne procède donc pas à des migrations de format. De même, les opérations de contrôle du contenu des versements est à la charge de l'établissement versant.

## 5. Faits marquants associés aux projets d'archivage électronique identifiés

### 5.1 Projets identifiés

Les entretiens avec les services d'archives ont permis d'identifier différents projets d'archivage électronique :

	<b>Service producteur</b>	<b>Objet</b>	<b>Type</b>	<b>Avancement projet</b>
AD31	Agence de l'eau de l'Adour	Dossiers d'études	Documents bureautiques	A l'étude
AD31	Laboratoire régional des ponts et chaussées	Etudes techniques	Documents bureautiques	A l'étude
AD35	Ministère du travail / Direction du travail	Conventions collectives / accords d'entreprises	Documents numérisés	A l'étude
AD35	DRAC	Recensement des monuments et sites à valeur patrimoniale	Données	A l'étude
AD42	Conseil Général	ANIS : Base de données sociales	Données	A l'étude
AD42	Conseil Général	Application gestion du personnel	Données	A l'étude
AD42	Conseil Général	Application gestion des infrastructures / Acquisitions de parcelles	Données	A l'étude
AD42	Conseil Général	Application gestion d'archives	Données	A l'étude
AD59	Préfecture du Nord	Dossiers de demandes de titres de séjour	Documents numérisés	Versement effectué
AD59	Préfecture du Nord	Dossiers d'associations	Documents numérisés	A l'étude
AD75	Ville de Paris / Direction de l'Urbanisme	Permis de construire	Données	Versement effectué
AD75	Ville de Paris	Gestion des courriers	Données	A l'étude
AD75	Ville de Paris	Liste électorale	Données	Versement effectué
AD75	Ville de Paris	Archives cabinet d'élu	Documents bureautiques	Versement effectué
AD78	Conseil Général	Application budgétaire en fin de vie	Données	A l'étude
AD78	Mairies, conseil général	Contrôle de légalité	Télé-procédure	A l'étude
AD78	Parquet de Versailles	Enregistrements / chaîne pénale	Données	Versement effectué
AD78	Conseil Général	Marchés publics	Télé-procédure	A l'étude
AD78	Conseil Général	Allocations personnalisées d'autonomie	Données	A l'étude
AD78	Etablissement public d'aménagement de la ville nouvelle de Saint Quentin	Sommier foncier électronique : application en fin de vie	Données	A l'étude

en Yvelines

AM Lyon	Ville de Lyon / Direction des Marchés	Marchés publics	Télé-procédure	A l'étude
AM Lyon	Ville de Lyon / Direction Finances	Application Finances en fin de vie	Données	A l'étude
AM Lyon	Ville de Lyon	Application délibérations en cours de mise en place	Documents bureautiques	A l'étude
AM Lyon	Ville de Lyon / Direction de l'urbanisme	Dossiers du service du logement	Documents numérisés	A l'étude
AM Lyon	Ville de Lyon / Mission organisation	Dossiers de missions	Documents bureautiques	Versement effectué
CAC	INSEE	Fichiers recensement	Données	Versement effectué
CAC	INSEE	Formulaires recensement	Documents numérisés	Versement effectué

On dénombre ainsi :

- 19 projets à l'étude
- 8 projets ayant déjà conduit à un versement
- 5 projets portant sur l'archivage de documents bureautiques
- 5 projets portant sur l'archivage de documents numérisés
- 14 projets portant sur l'archivage de données
- 3 projets portant sur l'archivage de documents échangés par télé-procédure

## 5.2 Catégorisation des projets identifiés

Le recensement des projets réalisés et en cours permet de distinguer trois grandes catégories de projets, suivant la nature des informations versées :

### 5.2.1 L'archivage de documents numérisés par le service producteur

#### Exemples types de projets :

Dossiers étrangers en préfecture (AD59), Déclarations d'associations (AD59), Conventions et accords d'entreprise de la direction du travail (AD35)

#### Principe :

Dans le cadre de la gestion courante de leurs dossiers physiques, certains services producteurs procèdent à leur dématérialisation en les numérisant puis en les conservant sous leur forme électronique.

Il s'agit essentiellement de dossiers. Ils se composent :

- De documents reçus par l'administration et signés par leur émetteur

- De pièces justificatives fournies par l'émetteur
- De documents produits par le service puis transmis après avoir été signés
- De documents complémentaires édités à partir des applications informatiques du service

Les documents ainsi numérisés sont conservés soit dans une arborescence de répertoires (poste de travail ou serveur), soit dans un logiciel de gestion électronique de documents après avoir fait l'objet d'une indexation.

#### **Motivations des services producteurs :**

Les services producteurs sont motivés par ce type de démarche principalement pour 3 raisons :

- Réduction de l'espace de stockage nécessaire à la conservation des dossiers durant leur durée d'utilité administrative,
- Réduction du temps de traitement et/ou de recherche sur ces documents,
- Intégration de ces documents avec des données contenues dans d'autres applications.

#### **Caractéristiques des documents :**

Les documents numérisés se caractérisent par :

- Un contenu fixe qui n'évolue généralement pas après la numérisation
- Un format généralement unique pour l'ensemble des documents
- Une taille de fichier généralement comparable d'un document à un autre (même format, mêmes conditions de numérisation ...)
- Des volumes stockés généralement importants (1 Go de données correspondant par exemple à 20.000 pages de 50 ko chacune).

#### **Contexte et processus d'archivage :**

On distingue :

- Les dossiers ayant un intérêt uniquement administratif et pouvant être détruit au terme de leur durée d'utilité administrative
- Les dossiers ayant un intérêt uniquement historique au-delà de leur durée d'utilité administrative et devant être conservés en tout ou partie sans limitation de durée
- Les dossiers ayant à la fois un intérêt historique et un intérêt administratif au-delà de leur durée d'utilité administrative et devant être conservés en tout ou partie sans limitation de durée

L'instruction DITN/DPACI/RES/2005/001 du 14 janvier 2005, portant sur les modalités de délivrance du visa d'élimination des documents papier transférés sur support numérique ou micrographique, prévoit que des visas de destruction puissent être accordés si les documents originaux sont éliminables au terme de leur DUA.

Dans le cas où les documents originaux ne peuvent pas être détruits, leur numérisation peut néanmoins présenter un intérêt pour le service producteur. Dans ce cas, l'usage des originaux devenant marginal durant la durée d'utilité administrative, il est à craindre que le service producteur ne procède plus au classement effectif des documents originaux, risquant ainsi de rendre leur utilisation, une fois ceux-ci archivés, très difficile.

Certains des documents concernés par cette numérisation peuvent présenter un caractère de preuve. La version électronique de ces documents peut sous certaines conditions garder cette valeur. Notamment, l'absence de mécanismes de signature électronique dès la phase de numérisation ou l'utilisation de supports non WORM peut compromettre cette valeur probante.

Deux scénarios d'archivage sont à prendre en compte :

- Le dossier est complet dès la phase de numérisation : les documents étant « fixes » dès leur création, l'opération d'archivage peut avoir lieu à tout moment y compris dès la numérisation
- Le dossier est complété progressivement durant son instruction : il est conseillé d'attendre au moins la clôture du dossier pour procéder à l'archivage (la notion de « statut » du dossier est alors fondamentale pour pouvoir procéder à l'archivage).

## 5.2.2 L'archivage de documents électroniques natifs détenus par les services producteurs

### Exemples types de projets :

Archives d'un cabinet d'élu de la ville de Paris, Archives de la mission organisation de la ville de Lyon, marchés publics, délibérations (contrôle de légalité).

### Principe :

Les échanges d'informations étant de plus en plus fréquemment dématérialisés, certains documents n'ont plus qu'une existence électronique et ne sont plus jamais édités. Leur espace de conservation électronique constitue par conséquent la source la plus sûre pour retrouver l'ensemble des documents relatifs à un thème donné dans leur dernière version.

Il peut s'agir :

- Soit de documents directement produits par l'administration : fichiers bureautiques, documents techniques (plans) ...
- Soit de documents reçus nativement sous forme électronique : messages électroniques, documents transmis par télé-procédures ...

Ces documents peuvent dans certains cas être signés électroniquement auquel cas il est important de pouvoir conserver également cette signature, au moins pendant la durée où la valeur de preuve subsiste.

### Motivations des services producteurs :

Les services producteurs sont motivés par l'archivage électronique de ces documents pour les raisons suivantes :

- L'édition physique de l'ensemble de ces documents engendrerait des coûts importants et occuperait un espace de classement superflu,
- Certains de ces documents ne peuvent pas être facilement matérialisés, soit parce qu'ils nécessitent des moyens spécifiques (plans), soit parce que les imprimer conduirait à perdre certaines informations (CAO), soit encore parce qu'ils ont été signés électroniquement et constituent donc des originaux ayant valeur de preuve.

### Caractéristiques des documents :

Les documents électroniques natifs se caractérisent par :

- Un contenu susceptible d'être modifié régulièrement par son auteur
- Des formats généralement très différents et liés aux applications utilisés pour les produire
- Des tailles de fichier très variées fonction des applications utilisées et des sujets traités
- Des volumes stockés pouvant dans certains cas être très importants (plans, présentations Powerpoint, éditeurs graphiques ...),
- Des méta-données généralement inexistantes excepté lorsque ces documents sont gérés dans un logiciel de gestion électronique de documents, ce qui reste relativement rare.

Exception : les documents transmis par télé-procédure sont généralement dès leur transmission fournis avec des méta-données. Les formats de documents utilisés peuvent dans certains cas être contrôlés en amont.

#### **Contexte et processus d'archivage :**

L'archivage de ce type de documents peut s'avérer très difficile du fait notamment de la très grande diversité des formats utilisés (qui ne sont pas nécessairement pérennes dans le temps) et des modes de gestion de ces documents.

Un archivage électronique peut être envisagé si :

- Tous les documents concernés sont produits sous des formats identiques et considérés comme pérennes dans le temps,
- Ils peuvent être fournis au service d'archives accompagnés de méta-données pertinentes.

Seuls les documents gérés dans des applications de GED et ceux transmis par télé-procédures peuvent, sous certaines conditions, être candidats à l'archivage électronique.

### **5.2.3 L'archivage de données contenues dans des applications**

#### **Exemples types de projets :**

Nouvelle chaîne pénale, fichier du recensement, application permis de construire

#### **Principe :**

Les données sont des informations structurées contenues dans des applications et produites soit après saisie par un utilisateur, soit suite à des traitements informatiques.

Les applications gérant ces données peuvent être de différents types :

- Applications locales spécifiques à l'administration
- Applications locales utilisées par plusieurs administrations (applications développées au niveau national ou progiciels du marché)
- Applications nationales hébergées au niveau national

Plusieurs services d'archives ont procédé au recensement des applications utilisées par leurs services producteurs, notamment les archives départementales de Paris et d'Ille et Vilaine.

Le recensement réalisé par les archives d'Ille et Vilaine a permis d'identifier 251 applications réparties dans 23 services :

	Local	National	?	Total
<b>Conseil Général</b>	<b>70</b>	<b>13</b>		<b>83</b>
Direction Régionale des Affaires Maritimes		1		1
Pôle Routes Infrastructures et Environnement	5	1		6
Tous pôles	7			7
Laboratoires départementaux	4			4
Pôle Administration et Ressources des Services	12			12
Comité départemental du Tourisme	2			2
Service Communication Information Documentation	4			4
Pôle Education Jeunesse Culture et Sports	15			15
POINT 35	1			1
Pôle Action Sociale	17			17
Direction Départementale des Affaires Sanitaires et Sociales	3	11		14
<b>Préfecture département</b>	<b>19</b>	<b>59</b>	<b>2</b>	<b>80</b>
Service Départemental d'Incendie et de Secours	2	1	1	4
Sous-préfecture de Redon	5	3		8
Services de la Préfecture	8	14	1	23
Direction Départementale du Travail de l'Emploi et de la Formation Professionnelle	2	11		13
Direction Départementale de la Protection Judiciaire de la Jeunesse		4		4
Direction de la Jeunesse et des Sports	2	26		28
<b>Préfecture région</b>	<b>8</b>	<b>75</b>	<b>5</b>	<b>88</b>
Direction Régionale de l'Équipement		9	3	12
Direction Régionale des Douanes	1	34		35
Direction Régionale de l'Environnement	5	11	2	18
Direction Interrégionale chargée des Anciens Combattants	1	11		12
Direction Générale de l'Aviation Civile	1	9		10
Direction Régionale du Commerce Extérieur		1		1
<b>Total</b>	<b>97</b>	<b>147</b>	<b>7</b>	<b>251</b>

On constate que les applications locales représentent l'essentiel des applications en conseil général. Inversement, les services déconcentrés utilisent essentiellement des applications nationales.

#### Motivations des services producteurs :

Les services producteurs sont motivés par l'archivage électronique de ces données pour les raisons suivantes :

- Certaines des données jugées pertinentes à conserver ne peuvent pas faire l'objet d'éditions papier et l'application qui les héberge n'est pas pérenne (elle va être arrêtée ou elle n'est pas suffisamment sécurisée)
- L'édition systématique d'états « papier » issus de ces applications engendrerait des coûts importants et occuperait un espace de classement superflu,
- La recherche d'une donnée pertinente sur un état papier peut s'avérer fastidieuse
- La conservation des données sous forme électronique permettrait de les exploiter ultérieurement (par exemple pour produire un état qui n'avait pas été envisagé avant)
- Certaines de ces données pourraient servir d'instrument de recherche pour identifier facilement un dossier physique archivé.

En pratique, la majorité des projets d'archivage de données électroniques portent aujourd'hui sur les données susceptibles de faciliter la recherche dans les dossiers papier versés, dans l'objectif notamment de palier à l'absence de registres de plus en plus fréquente. Le demandeur est donc la plupart du temps le service d'archives et non le service producteur.

### **Caractéristiques des données à archiver :**

Dans la plupart des cas, les données ayant vocation à être archivées se matérialisent par une ou plusieurs fichiers correspondant à des tables. Néanmoins, pour que ces fichiers soient exploitables, ils doivent être complétés par d'autres fichiers ce qui peut conduire à plusieurs dizaines de tables au final.

Les données se caractérisent par :

- Un contenu structuré non directement intelligible. Un enregistrement, pour pouvoir être lu nécessite de connaître la signification des champs qui lui sont associés ainsi que les liens utilisés vers d'autres tables de données. Chaque application dispose de son propre modèle de données s'appuyant sur des méthodologies différentes et nécessitant des compétences techniques précises.
- Un format dépendant de l'application utilisatrice, souvent propriétaire à cette application
- Les volumes représentés par les données à archiver ne sont généralement pas très importants

### **Contexte et processus d'archivage :**

Deux cas peuvent se produire :

- Les données n'ont pas de vocation à un archivage historique. Dans ce cas, l'archivage des données dépend de la durée de vie de l'application et des conditions de préservation des données dans cette application :
  - o Si la durée de vie de l'application est supérieure à la durée d'utilité administrative des données et si l'application en garantit l'intégrité et la restitution, leur archivage en dehors de l'application source n'est pas nécessaire
  - o Sinon, un processus d'archivage électronique « intermédiaire » doit être mis en place. Il portera sur l'ensemble des données concernées par la DUA.
- Les données ont une vocation historique et doivent être conservées en totalité ou partiellement sans limitation de durée : dans ce cas, un processus d'archivage électronique doit être implémenté.

Certaines applications produisent des états électroniques présentant les données qu'il est nécessaire d'archiver sous un format exploitable pour l'archivage électronique.

Quand c'est le cas, l'archivage des données peut être remplacé par l'archivage de ces documents.

Quand ce n'est pas le cas, il faut soit faire évoluer le logiciel de façon à ce qu'il intègre une fonction d'édition de ces états électroniques, soit extraire les données pertinentes et les traiter de façon à ce que leur préservation soit assurée et qu'elles soient exploitables à terme par leurs futurs utilisateurs.

L'archivage électronique de données implique par conséquent deux types d'actions distinctes :

- Le paramétrage de l'application pour que celle-ci puisse éditer les états électroniques ou extraire les données pertinentes sous une forme compatible avec les objectifs de pérennité et de restitution de l'archivage électronique. Ce paramétrage est à réaliser une fois par application.
- Le versement des archives. Ce versement peut être unique (à la fin de vie d'une application) ou périodique.



La responsabilité de l'archivage peut dépendre d'une ou plusieurs administrations :

- Lorsque l'application est locale et spécifique à l'administration utilisatrice : les opérations de paramétrage et de versement sont à la charge de l'administration utilisatrice.
- Lorsque l'application est locale mais utilisée par plusieurs administrations (applications développées au niveau national ou progiciels du marché) : les opérations de paramétrage peuvent être effectuées une seule fois soit par l'éditeur, soit par le maître d'œuvre national de l'application. Les opérations de versement doivent par contre être effectuées par le service utilisateur.
- Lorsque l'application est nationale et hébergée au niveau national : les opérations de paramétrage ne peuvent généralement être faites que par le maître d'œuvre national, le service utilisateur ne disposant pas nécessairement des moyens suffisants pour procéder à des paramétrages ou développements complémentaires. Les versements peuvent quant à eux être effectués soit par le service responsable au niveau national de l'application, soit par le service utilisateur local. Néanmoins, l'opération sera vraisemblablement mieux supportée par le service national.

## 5.3 Sources d'informations candidates à l'archivage électronique

Pour qu'une source d'information puisse faire l'objet d'un archivage électronique, il est indispensable qu'elle réponde aux conditions suivantes :

- Les informations à archiver doivent être fournies dans un format pérenne c'est-à-dire ouvert et si possible non propriétaire,
- Elles doivent être fournies avec des méta-données permettant de les retrouver facilement.

Les sources d'informations répondant à ces conditions et pour lesquels un archivage électronique semble prioritaire sont de trois types :

- **Les fonds numérisés et documents électroniques natifs gérés dans des bases documentaires (GED)** : ces bases documentaires ont l'avantage d'associer à chaque document ou groupe de documents des méta-données et de traiter tous les documents relatifs à un thème d'information précis en organisant souvent les documents de façon similaire à la gestion des dossiers physiques. L'archivage du contenu de ces bases peut donc potentiellement remplacer l'archivage papier des dossiers.
- **Les documents transmis par télé-procédure** : ces documents ont la caractéristique d'être difficilement matérialisables du fait notamment de la signature électronique qui leur est associée. Ils ont la plupart du temps un caractère de preuve qu'il faut conserver. Bien que la part de ce type de documents soit encore faible, leur développement dans les prochaines années semble évident, à condition que des mécanismes d'archivage électronique fiables puissent leur être associés.
- **Les données administratives issues d'applications** : ces applications sont aujourd'hui très nombreuses dans les administrations mais l'archivage de leurs données est rarement réalisé, excepté par l'intermédiaire des états papier qui sont produits à partir de ces applications. Or ces états ne sont généralement pas conçus et produits dans une logique d'archivage à long terme et ne portent pas nécessairement sur les informations qu'il est pertinent de conserver. Ceci étant, des opérations d'archivage électronique ont déjà été réalisées (et de nombreux projets sont en cours), le principe étant de procéder à l'extraction de données à partir de

quelques tables intéressant les services d'archives (extraction au format ASCII délimité ou XML suivant la structure de la base de données).

## 5.4 Projets types

Parmi l'ensemble des projets identifiés, nous avons sélectionné 4 projets représentatifs des contraintes liées à l'archivage des différents types de sources candidats à l'archivage électronique. Ces projets ont déjà conduit à des versements ou ont déjà fait l'objet d'une spécification.

### 5.4.1 Bureau des nationalités de la préfecture du Nord : numérisation des dossiers de demandes de carte de séjour (AD59)

Le service des archives du Nord a été impliqué dans un projet de mise en place d'un système de gestion électronique de documents relatif à la gestion des dossiers de demandes de titres de séjour (18.000 dossiers ouverts chaque année, 110.000 dossiers considérés comme actifs). Ces dossiers se composent d'un formulaire CERFA de demande de titre et de pièces justificatives fournies par le demandeur.

Les objectifs de ce projet étaient de :

- faciliter les opérations de recherche d'informations par les agents du bureau
- limiter l'espace de stockage consacré aux dossiers papier (le bureau des nationalités détenait au lancement du projet 2.500 mètres linéaires de dossiers).
- Préserver la documentation historique de la recherche.

Le système mis en place dans le service producteur permet de consulter l'ensemble des documents reçus de la part des demandeurs de titres de séjour.

Une fois un dossier déposé, celui-ci est entièrement numérisé par un service spécialisé. Les documents numériques ainsi produits sont ensuite enregistrés et indexés dans la GED, elle-même interfacée avec le logiciel national de gestion des ressortissants étrangers AGDREF.

Les dossiers papier ne sont, par conséquent, plus conservés qu'une seule année puis détruits à l'exception des dossiers soumis au contentieux et des formulaires CERFA originaux qui restent archivés sous forme papier du fait de la présence de la signature du demandeur sur ces documents.

Le versement s'effectue selon les étapes suivantes :

- Au niveau du service producteur :
  - o Extraction des données et documents dont la DUA est dépassée (10 ans),
  - o Classement des documents et sélection des documents à conserver (environ 10%)
  - o Gravage d'un CD-Rom (Worm) conforme à la norme AFNOR NF Z 42-011-2
- Au niveau du service d'archives :
  - o Contrôle de la qualité du gravage
  - o Nommage des CD-Rom

- Rédaction du bordereau de versement incluant les méta-données associées

Les documents versés sont des documents PNG accompagnés de fichiers aux formats TXT et XML pour les méta-données.

En parallèle, un protocole de versement des instruments de recherche associés doit être défini.

La recherche s'effectue dans un premier temps à l'aide des informations contenues dans le bordereau de versement à partir du logiciel CLARA.

Une fois le CD-ROM contenant les documents recherchés identifié, la recherche des documents sur le support s'effectue à l'aide d'un programme auto-exécutable depuis le CD-ROM et permettant de consulter le fichier XML sur Internet Explorer.

La prise en compte de l'archivage électronique dans le projet a consisté à :

- Préciser quelles étaient les informations pour lesquelles l'archivage électronique était envisageable,
- Préciser les règles à respecter au niveau de l'application de GED pour garantir la présomption d'authenticité des documents,
- Définir le format associé aux méta-données
- Définir les règles de sélection à appliquer pour l'archivage définitif
- Rédiger un nouveau tableau de gestion tenant compte des nouvelles règles de conservation

Ces documents ne seront pas communicables avant 60 ans, excepté si des travaux de recherche spécifiques sont prévus et qu'une dérogation est accordée.

#### **Données quantitatives disponibles :**

Estimation annuelle du nombre de dossiers archivés : 270

Nombre moyen de pages par dossier : 25 pages

Taille moyenne des pages : 50 ko

Taille moyenne d'un dossier : 1,25 Mo

Volume moyen versé par an pour le département du Nord : 337 Mo

Volume archivé au bout de 10 ans : 3,37 Go

### **5.4.2 Conseil général des Yvelines : Délibérations et contrôle de légalité (AD78)**

Le conseil général des Yvelines expérimente depuis plusieurs mois la plate-forme de dématérialisation des échanges FAST proposée par la Caisse des Dépôts et Consignations.

L'objet de l'expérimentation porte sur les opérations de contrôle de légalité effectuées entre le conseil général et la préfecture : les délibérations du conseil général et les arrêtés des ressources humaines doivent être transmis à la préfecture (entre 100 et 200 décisions par mois).

Le système FAST permet au conseil général de télé-transmettre ses décisions à la préfecture sous la forme de documents PDF signés et de recevoir en échange un accusé de réception de la part de la préfecture.

Les documents en question n'ont donc plus besoin d'être imprimés pour être transmis. La version originale du document est celle stockée dans la plate-forme avec la signature.

Les documents peuvent toujours être imprimés mais ils ne peuvent plus avoir d'existence physique au sens légal (preuve) puisque la signature électronique ne peut pas être matérialisée.

Or il est de la responsabilité de la collectivité territoriale de les conserver durant la DUA avant de les transmettre au service d'archives compétent pour l'archivage définitif.

La direction des Archives de France en collaboration avec la Caisse des Dépôts et Consignations a modélisé les transactions « versement » et « restitution » à partir de FAST vers une plate-forme d'archivage, l'Archive se composant de l'objet Archives et de sa description structurée suivant la DTD EAD et encapsulant les documents stockés sur la plate-forme.

#### **Données quantitatives disponibles :**

Nombre de transmissions réalisées par le conseil général des Yvelines du 1<sup>er</sup> janvier 2004 au 28 février 2005 : 1920

Nombre de documents métiers transmis : 1930 (soit 1,01 documents par transmission)

Taille moyenne des documents transmis : 80 ko

Taille moyenne des transmissions (en-tête compris) : 901 ko

Evaluation du nombre annuel de documents émis par an par le conseil général (arrêtés des ressources humaines et délibérations du conseil général et commissions permanentes) : 2800 (soit 250 documents par mois).

Evaluation du volume annuel de documents à verser : 2.523 Mo.

Evaluation du volume des données transmises par an pour le Conseil Général : 1,5 Go

Si ces données étaient archivées en totalité, elles représenteraient donc au bout de 10 ans un volume de 15 Go de données.

### **5.4.3 Juridictions pénales : la nouvelle chaîne pénale (NCP)**

Le ministère de la Justice a développé et déployé dans différents tribunaux de grande instance, un logiciel de suivi de la chaîne Pénale permettant notamment l'enregistrement et le suivi des procédures pénales.

L'enregistrement réalisé pour chaque procédure porte notamment sur la nature de l'affaire, la date de début des faits, le type d'infraction ...

Ce logiciel comprend un module d'archivage qui prévoit d'extraire les données relatives aux affaires terminées depuis 10 ans (DUA). Ces données étaient jusque-là communiquées aux services d'archives sous la forme de registres.

Chaque année, les juridictions utilisent le module d'archivage pour extraire les données relatives aux dossiers dont la DUA est désormais dépassée. Cette extraction prend la forme de fichiers TXT représentant certaines tables de l'application. Une dizaine de tables sont ainsi extraites dont notamment deux tables principales : une table « Affaire » et une table « Personne ». Ces deux tables sont liées par un champ « Numéro d'affaire » présent dans la table « Personne » et renvoyant à l'affaire correspondante dans la table « Affaire ». D'autres tables annexes permettent de lier les codes fournis au niveau des tables « Affaire » et « Personne » avec leurs libellés effectifs.

Pour l'instant, les extractions portent uniquement sur les affaires sans suite dont les dossiers ne sont pas archivés. Les extractions pourraient également porter sur les affaires dont les dossiers sont archivés, auquel cas ces données permettraient de faciliter la recherche des dossiers et seraient utilisées comme instrument de recherche.

Les premiers versements ont eu lieu en 2004. Le versement a consisté à inscrire les tables extraites sur un CD-Rom puis à transmettre ce CD-Rom au service d'archives concerné.

Une application de consultation des données archivées a été développée.

Le service d'archives de Paris a dû, pour pouvoir exploiter ces données, les importer dans une base Access puis reconstruire les liens entre les tables. En parallèle, le Ministère de la Justice a réalisé un développement complémentaire permettant de faciliter leur consultation.

#### **Données quantitatives disponibles :**

Le versement réalisé à Paris en 2004 correspondait à 900.000 enregistrements représentant un volume de 250 Mo (soit 0,28 Ko par enregistrement). Il correspond aux affaires sans suite des années 73 à 93 (soit près de 43.000 enregistrements par an).

Si on fait l'hypothèse que 50.000 enregistrements seront versés chaque année pour Paris, le volume versé serait par conséquent d'environ 13,9 Mo.

Evaluation du volume archivé au bout de 10 ans : 139 Mo

#### **5.4.4 Direction de l'urbanisme de la mairie de Paris : données relatives aux permis de construire (AD75)**

L'application GERCO de la direction de l'urbanisme de la mairie de Paris gère les données relatives aux dossiers de permis de construire instruits par la mairie.

Le service des archives de Paris s'est montré intéressé par une partie des données gérées par cette application car plus de 7000 consultations sont réalisées chaque année en salle de lecture sur les dossiers papiers correspondant. Ce fonds représente aujourd'hui 2 km linéaires au sein du service et leur communication au public est immédiate dès leur versement à condition que le dossier soit clos.

Cette application gérant les données utiles pour identifier et retrouver facilement un dossier, celles-ci peuvent donc servir d'instrument de recherche en salle d'archives.

Deux versements ont été effectués à ce jour. Le rythme de versement devrait être annuel et porter sur l'ensemble des dossiers considérés comme clos depuis plus de 5 ans (DUA) dans GERCO.

Les données sont exportées sous format texte puis converties au format XML avant d'être livrées par CD-Rom au service d'archives.

Le service des archives et la direction de l'urbanisme ont travaillé ensemble pour spécifier les conditions d'archivage de ces données. Les étapes de cette spécification ont été les suivantes :

- Spécification des données à récupérer et des dates de versement : 2 à 3 réunions de travail.  
Règle définie : Les données doivent être versées dès qu'est renseigné le numéro de carton correspondant
- Extraction des données de l'application GERCO (service producteur) : exportation des données sous forme de tables
- Passation d'un marché portant sur la réalisation d'un outil de conversion des données au format XML et sur la reprise des données extraites (30.000 €)
- Définition d'une grille de conversion des données vers le modèle EAD (XML)
- Développement de l'application de conversion
- Reprise des données
- Réalisation d'une base Access de consultation

**Données quantitatives disponibles :**

Les données relatives à 8000 permis aboutis représentent 8 Mo de données brutes et 17 Mo de données après conversion XML.

Estimation du volume archivé au bout de 10 ans : 170 Mo