

Direction des archives de France
Cahier des charges d'une étude
sur la réalisation de plates-formes d'archivage électronique
dans les services publics d'archives

1. Le contexte

1.1. Environnement juridique

Les archives sont l'ensemble des documents, quels que soient leur date, leur forme et leur support matériel, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité (article L 211-1 du Code du Patrimoine) et leur conservation est organisée dans l'intérêt public tant pour les besoins de la gestion et de la justification des droits des personnes physiques ou morales, publiques ou privées, que pour la documentation historique de la recherche (article L 211-2 du Code du Patrimoine).

Dans ce cadre, les archives publiques (notamment les archives des services de l'Etat, ainsi que celles provenant des collectivités territoriales)¹ sont gérées, à un moment donné de leur cycle de vie², par les Archives nationales ou par le réseau des archives départementales et municipales³, ou régionales. Les services d'archives régionales, départementales et municipales sont des services décentralisés (article L 212-8 du code du Patrimoine) mais un contrôle technique et scientifique de l'Etat s'exerce par l'intermédiaire des directeurs d'archives départementales qui sont des personnels de l'Etat (article L 212-10 du code du Patrimoine).

Tous les documents produits ne sont pas archivés : à l'issue de leur durée d'utilité administrative, ils font l'objet, en accord entre l'administration des archives et le service producteur, d'un tri visant à sélectionner les documents destinés à être éliminés (avec le visa du service d'archives) et ceux destinés à être versés au service d'archives (article L 212-3 du code du Patrimoine).

¹ Les archives publiques sont :

- a) Les documents qui procèdent de l'activité de l'Etat, des collectivités territoriales, des établissements et entreprises publics ;
- b) Les documents qui procèdent de l'activité des organismes de droit privé chargés de la gestion des services publics ou d'une mission de service public ;
- c) Les minutes et répertoires des officiers publics ou ministériels (article L 211-4 du code du Patrimoine).

² On distingue trois étapes dans la vie d'une archive : archive courante (l'archive est encore d'utilisation courante) ; archive intermédiaire : elle n'est plus d'utilisation courante mais elle doit être conservée en raison de sa valeur probante (durée d'utilité administrative) ; archive définitive : l'archive, une fois dépassée la durée d'utilité administrative, est conservée en raison de sa valeur patrimoniale. Elle est à ce stade là conservée au sein du service public d'archives. Durant la période intermédiaire, qui varie beaucoup suivant la catégorie de documents (de quelques mois à 100 ans), la gestion directe se fait soit par le producteur soit par le service public d'archives.

³ Toutes les municipalités n'ont pas un service d'archives constitué. Les communes dépourvues d'un tel service peuvent confier leurs archives aux services d'archives départementales (obligation pour les archives de plus de 30 à 150 ans, suivant leur nature, des communes de moins de 2000 habitants : article L 212-11 du code du Patrimoine).

1.2. Informatisation et archivage électronique dans les services d'archives aujourd'hui

1.2.1. L'informatisation

Les services d'archives ont largement informatisé leur propre gestion en développant ou acquérant des logiciels spécialisés⁴, permettant de gérer les entrées des documents, le récolement⁵, la communication des archives aux lecteurs...

Ces logiciels possèdent également des modules documentaires permettant de décrire la nature des archives versées et de les retrouver à partir de plusieurs clés d'accès (dont des index, généralement à partir de dictionnaires de termes hiérarchisés). Ces modules documentaires sont moins largement répandus que les modules de gestion. Certains services ont particulièrement mis l'accent sur cet aspect documentaire et ont mis en place par ailleurs un accès en ligne aux inventaires sur Internet, comme par exemple les archives départementales du Puy de Dôme (<http://www.puydedome.com/accueil.htm>), de la Loire (<http://www.loire.fr/html/fr/archives/pages/accinv.htm>) ou de la Savoie http://www.sabaudia.org/v2/archives73/series/73_inventaire.php.

Parallèlement une norme internationale de description des documents d'archives (ISAD-G) est parue, qui a fait l'objet d'une DTD (definition type document) XML : EAD (Encoded Archival Description), version 2002 (<http://www.archivesdefrance.culture.gouv.fr/fr/archivistique/index.html>).

Les services d'archives ont par conséquent commencé à utiliser cette DTD, soit par l'encodage direct d'inventaires en EAD, au moyen d'éditeurs, soit parce que certains des logiciels de description des archives évoqués ci-dessus⁶ permettent de transformer automatiquement en EAD les données descriptives saisies dans des bases de données. Des applicatifs ont également commencé à être développés permettant la publication et la recherche de ces inventaires sur le réseau Internet : on peut citer, comme outil libre de droits, la plate-forme SDX de publication de documents XML (<http://adx.org/sdx/>) et son applicatif de publication et recherche d'instances EAD Pleade (<http://www.pleade.org/fr/>).

En outre, les services d'archives sont nombreux à numériser les fonds⁷ le plus fréquemment consultés (ou particulièrement précieux et/ou fragiles) : registres paroissiaux, registres d'état civil, plans cadastraux, registres matricules.... Quatre services d'archives départementales sont aujourd'hui à la pointe de ce mouvement (en terme de volume de données disponibles accessibles par réseau) : les archives départementales de la Mayenne (<http://perso.wanadoo.fr/archives.53/>) ; celles de la Savoie, de la Vendée (<http://archives.vendee.fr/>) ou encore des Yvelines (<http://www.cg78.fr/archives/>). Dans ce dernier service, plus de 3 millions d'actes sont ainsi numérisés, stockés sur des serveurs hébergés par les Archives départementales suivant une architecture SAN permettant d'abriter de 4 à 10 TO. Les documents numérisés disponibles sur Internet sont en général accessibles gratuitement, mais parfois contre paiement (cas de la Savoie). Pour cette diffusion, des applications spécifiques ont été développées sans que des liens aient encore été effectuées

⁴ Cinq à six logiciels sont utilisés par la quasi totalité des services d'archives régionales, départementales et municipales : Arkhéia, Avenio, Clara, Gaïa, Thot...

⁵ Localisation des archives (conditionnées généralement dans des boîtes, cartons, fichiers, liasses, registres) sur les rayonnages implantés dans les magasins d'archives.

⁶ Les plus avancés sont Arkhéia <http://www.arkheia.net/SOCIETE/PRESENTATION> et Gaïa <http://www.gaia-archives.org/public/Histo.htm>

⁷ Soit à partir de microfilms, soit à partir des originaux papier.

avec le système d'information général des services d'archives (logiciels de gestion et de description des archives évoqués ci-dessus).

On peut également citer un service d'archives municipales, à savoir les archives municipales de Rennes qui, d'une part, publie sur internet leurs inventaires (à partir d'extractions réalisées par le logiciel Avenio) et, d'autre part, ont associé à certains de ces inventaires, 300 000 images (<http://www.archives.rennes.fr/fonds>).

D'autres services ont également numérisé d'importantes collections mais qui ne sont accessibles qu'en local.

Les Archives nationales (centre historique des archives nationales et centre des archives d'outre-mer notamment) se sont également lancées dans de massives opérations de numérisation : on peut citer le cas de la numérisation des microfilms des archives de l'état civil des européens d'Algérie, réalisée dans le cadre de l'année de l'Algérie, soit plus d'un million d'actes prochainement consultables sur Internet⁸, à partir d'un autre applicatif développé avec SDX : Navimages (<http://www.archivesnationales.culture.gouv.fr/caom/Djazair/pages/appli/appli.html> et <http://sdx.archivesdefrance.culture.gouv.fr/gpl/navimages/index.html>).

1.2.2. L'archivage électronique

Les principes concernant la gestion et l'organisation des archives publiques en France restent identiques pour les archives électroniques.

Même si l'administration a largement informatisé ses procédures depuis plusieurs décennies, la valeur probante des documents a continué à résider dans les documents papier (signés manuscritement et accompagnés de nombreux autres signes de validation). Il s'ensuit que le papier a continué à largement dominer dans la sphère administrative : actes législatifs et réglementaires, dossiers individuels ou de principe auxquels on accède par des enregistrements informatisés. Durant ces dernières années, on observe une utilisation massive de la bureautique (et par conséquent une production croissante de fichiers électroniques : traitement de texte et tableurs essentiellement) ainsi que de la messagerie. De même, la généralisation des réseaux intranet/extranet et Internet posent la question de la conservation de leur contenu. Ceci étant, sauf à titre expérimental, les services d'archives en France n'ont pas encore eu à traiter ce type d'informations, d'autant que les pratiques de records management, permettant l'enregistrement de toute information entrante et sortante d'une institution ainsi que sa traçabilité, ne sont pas encore appliquées, sauf exception, en France⁹.

Par conséquent, les services d'archives n'ont pratiquement pas eu à recevoir des archives électroniques, à l'exception notable du Centre des archives contemporaines (service Constance) (<http://www.archivesnationales.culture.gouv.fr/cac/fr/index.html> et <http://www.archivesdefrance.culture.gouv.fr/en/archives%20administration/DAFcongresconstance.htm>) qui, dès le début des années 1970, a reçu des fichiers statistiques¹⁰ (INSEE, INED, recensement de la population...) et a développé, dans ce cadre, une méthodologie concernant la collecte (constitution des métadonnées) et le stockage (politique de migrations successives de formats, de supports). Le service de restitution des données archivées (excepté aux services

⁸ Actuellement la base de données est consultable sur Internet mais les actes eux-mêmes doivent être consultés sur place.

⁹ Malgré la sortie d'une norme ISO 15489 sur le sujet : [http://www.archimag.com/articles/148/148records1\(norme\).html](http://www.archimag.com/articles/148/148records1(norme).html).

¹⁰ Plus de 6000 (300 GO). Un versement massif de plusieurs TO est attendu avec le dernier recensement de la population.

versants) n'a pas été développé, les fichiers statistiques n'étant pas communicables avant un délai de 100 ans. Les versements électroniques de Constance sont enregistrés dans le système d'information du centre : Priam 3 (<http://www.archivesnationales.culture.gouv.fr/cac/fr/index.html>).

Les archives départementales ont commencé, dans les trois dernières années, à recevoir quelques versements électroniques, notamment de bases de données (bases de données provenant de la Justice, permis de construire) comme aux Archives de Paris : <http://www.archives.paris>.

De leur côté, les archives départementales du Nord ont travaillé avec les services de la préfecture pour l'intégration dans une application de GED concernant la gestion et le suivi des dossiers de titres de séjour, de la fonction archivage avec la prise en compte du cycle de vie des documents et l'intégration de fonctionnalités permettant d'extraire, à intervalles réguliers, sur CD-R, des dossiers pour les Archives départementales (vdoom@cg59.fr).

Par ailleurs, les archives départementales de l'Ile et Vilaine ont confié à une société une étude sur les scénarios qui pourraient être envisagés pour mettre en place l'archivage électronique pour les services de l'Etat dans le département. Des visites de plusieurs services ont été réalisées, permettant de faire un état des lieux de la production électronique (dans la très grande majorité des cas, des bases de données) et, à partir de là, de définir des priorités et d'envisager trois types de modèles organisationnels suivant que les archives départementales jouent également le rôle de service d'archives intermédiaires ou non¹¹. Enfin, l'étude fait le choix de versements et de conservation sur des supports optiques type CD-R. Le rapport est disponible aux archives départementales d'Ile et Vilaine (jean-yves.le-clerc@cg35.fr).

En revanche, certaines grandes institutions scientifiques qui produisent et reçoivent, depuis plusieurs dizaines d'années, des données uniquement électroniques, ont déjà organisé des services d'archivage électronique sur toute la chaîne de traitement, et notamment le CNES à Toulouse. Trois services existent, coordonnés entre eux : le service « collecte et préparation des données », le service « stockage d'archives » et le service « gestion des données et communication ». Le service « stockage », ou STAF, est en place depuis 1994, son objectif étant de ranger des « collections » de fichiers selon un logique applicative et sans se soucier des évolutions des systèmes et technologies. Actuellement, le STAF a la responsabilité du stockage intègre de 3,8 millions de fichiers pour un volume de 145 TO (claude.huc@cnes.fr). Les options choisies par le CNES pourraient être appliquées aux services d'archives publics.

1.3. Le développement de l'e-administration et ses conséquences

1.3.1. Le contexte

L'archivage électronique est amené à se développer très rapidement dans les années à venir (horizon 2005-2010) dans le cadre du développement de l'e-administration (programme ADELE, voir http://www.adae.gouv.fr/rubrique.php3?id_rubrique=3). En effet, la plupart des informations circulant entre administrations, entre les citoyens et l'administration, seront progressivement dématérialisées sous la forme de flux de données transitant par des espaces sécurisés. Les informations électroniques auront valeur probante, évolution rendue possible

¹¹ En recevant des documents n'ayant pas encore atteint la fin du délai d'utilité administrative.

par la nouvelle législation parue en la matière depuis 3 à 4 années, accordant la même valeur aux documents sur support papier et sur support électronique, à condition que ces derniers soient conservés dans des conditions permettant d'assurer leur intégrité : loi 2000-230 du 13 mars 2000 sur l'adaptation du droit de la preuve aux nouvelles technologies et à la signature électronique (et ses décrets d'application), loi 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique. Dans le cadre de ces flux, les supports papier ne devraient plus être conservés (utilisation en tant que documents de travail, brouillons...), sauf en cas de re-matérialisation à un moment donné de la chaîne ou en fin de processus.

Généralement l'objet télétransmis s'accompagne d'un formulaire (structuré en XML et défini par des normes métier) permettant de caractériser le flux et par conséquent de le retrouver.

Parmi les productions de ce nouveau type figurent des projets comme la dématérialisation du contrôle de légalité ou celle des marchés publics, rendue obligatoire par l'article 56 du code des marchés publics (annexé au décret n° 2004-15 du 7 janvier 2004 portant CMP). D'autres projets sont en cours comme la dématérialisation des demandes d'extraits de l'état civil, ou encore celle des dossiers de RMI-RMA, dématérialisation de la comptabilité publique...

Ce nouveau type de dématérialisation implique la mise en œuvre d'une plate-forme sécurisée d'échanges, comme celle développée actuellement par la Caisse des dépôts et consignations (FAST, voir <http://www.dent.caissedesdepots.fr/commun/pdf/publications/horizon/hn1-page2.pdf>), qui permet de réunir sur une même plate-forme les services de tiers de confiance (pour les scellements et signatures), de tiers horodateurs... FAST est actuellement utilisé pour tester la dématérialisation du contrôle de légalité¹².

Les différentes plates-formes d'échanges qui pourraient être mises en place sur ce modèle n'ont pas vocation à assurer l'archivage des données. Par conséquent, les informations qui doivent être conservées, soit pour faire preuve soit en raison de leur intérêt patrimonial, doivent être transférées sur des plates-formes d'archivage, dans les services publics d'archives (pour ce qui concerne les documents à conserver définitivement) : Archives nationales (service Constance puis, à l'horizon 2010, dans le nouveau centre des archives nationales programmé à Pierrefitte), archives départementales, voire archives régionales et municipales¹³.

Au-delà du versement, ces informations devront être conservées par les services d'archives pour restitution aux services producteurs, lorsque ceux-ci souhaiteront consulter une nouvelle fois leurs données, ainsi qu'à l'ensemble des citoyens à l'expiration des délais de communicabilité¹⁴.

¹² Les collectivités territoriales envoient aux préfetures leurs actes réglementaires. Si aucune observation n'est émise par les préfetures au terme d'un délai de 2 mois, les actes sont réputés valides. Les actes entre les collectivités et les préfetures transitent par FAST, suivant la norme métier ACTES, répertorié sur le site de l'ADAE (répertoire des schémas XML : http://www.adae.gouv.fr/rubrique.php3?id_rubrique=43).

¹³ Pour les plus gros services existants.

¹⁴ Article L213-1 du Code du Patrimoine : "Les documents dont la communication était libre avant leur dépôt aux archives publiques continueront d'être communiqués sans restriction d'aucune sorte à toute personne qui en fera la demande.

Les documents mentionnés à l'article 1er de la loi n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal demeurent communicables dans les conditions fixées par cette loi.

1.3.2. L'expérimentation avec la CDC et le conseil général des Yvelines

Sur la base de contacts établis entre la Caisse des dépôts et consignations et la Direction des Archives de France, un document sur l'articulation entre FAST et une plate-forme d'archivage a été rédigé par la CDC et deux projets de normes (schémas XML) ont été réalisés en collaboration entre les deux institutions.

L'un vise à modéliser les échanges entre service versant et service d'archives autour de deux transactions principales : le versement (versement / notification en synchrone / accusé de réception en asynchrone ou anomalie) et la restitution (demande / notification / restitution ou anomalie / accusé de réception de la restitution ou anomalie). Cette modélisation est générique et vise à couvrir tout type de versement, quelle que soit la nature des informations archivées.

L'autre vise à modéliser la description de l'archive ainsi transférée, l'archive étant constituée de l'objet même à conserver et de sa description. Cette description varie suivant les processus métier dématérialisés mais **doit obéir à la DTD EAD** (qui permet d'avoir un inventaire des données transférées) évoquée ci-dessus, dans le cadre du standard METS (metadata encoding and transmission standard : <http://www.loc.gov/standards/mets/>), qui est une implémentation de l'OAIS. Ainsi la plate-forme d'échanges doit permettre de transférer les données sur la plate-forme d'archivage suivant la norme de versement avec parallèlement la génération de fichiers METS pour la description des informations transférées.

Par ailleurs, le conseil général des Yvelines s'est porté candidat pour évaluer les coûts de la mise en œuvre de la plate-forme d'archivage permettant à sa direction des archives départementales de recevoir les informations provenant de FAST et concernant différentes collectivités territoriales des Yvelines qui se sont portées volontaires pour tester la dématérialisation du contrôle de légalité, étant entendu que le service informatique des Yvelines a déjà mis en œuvre au sein des archives départementales une infrastructure de stockage (SAN) ainsi qu'une infrastructure et des applicatifs permettant aux internautes d'accéder aux ressources électroniques provenant de la numérisation de certains fonds des archives (voir 1.2.1 ci-dessus).

Tous les autres documents d'archives publiques pourront être librement consultés à l'expiration d'un délai de trente ans ou des délais spéciaux prévus à l'article L 213-2."

Article L 213-2 du Code du Patrimoine :

"Le délai au-delà duquel les documents d'archives publiques peuvent être librement consultés est porté à :

a) Cent cinquante ans à compter de la date de naissance pour les documents comportant des renseignements individuels de caractère médical ;

b) Cent vingt ans à compter de la date de naissance pour les dossiers de personnel ;

c) Cent ans à compter de la date de l'acte ou de la clôture du dossier pour les documents relatifs aux affaires portées devant les juridictions, y compris les décisions de grâce, pour les minutes et répertoires des notaires ainsi que pour les registres de l'état civil et de l'enregistrement ;

d) Cent ans à compter de la date de recensement ou de l'enquête, pour les documents contenant des renseignements individuels ayant trait à la vie personnelle et familiale et, d'une manière générale, aux faits et comportements d'ordre privé, collectés dans le cadre des enquêtes statistiques des services publics ;

e) Soixante ans à compter de la date de l'acte pour les documents qui contiennent des informations mettant en cause la vie privée ou intéressant la sûreté de l'Etat ou la défense nationale et dont la liste est fixée par décret en Conseil d'Etat."

2. L'étude à réaliser

En complément de l'étude ponctuelle qui pourrait être menée par le Conseil général des Yvelines, la Direction des archives de France (Francoise.Banat-Berger@culture.gouv.fr ; olivier.de-solan@culture.gouv.fr) souhaite faire réaliser une étude plus générale sur le coût de mise en œuvre de plates-formes d'archivage électronique (coûts humains, matériels, logiciels...), d'une part au sein du service Constance des Archives nationales, d'autre part dans les services d'archives départementales, voire dans certains gros services d'archives municipales ou dans les services d'archives régionales.

Cette évaluation prendra la forme d'un plan projet de mise en œuvre, incluant les plannings, les charges et les coûts détaillés.

Cette évaluation devra également présenter les divers pré-requis nécessaires à la mise en œuvre et à la gestion d'une telle infrastructure sur le plan humain (compétences nécessaires) et matériel (locaux – machines) notamment.

2.1. Fonctionnalités attendues de la plate-forme d'archivage

Celles-ci sont basées sur le modèle OAIS (norme ISO 14721:2003 (Systèmes de transfert des informations et données spatiales -- Système ouvert d'archivage de l'information -- Modèle de référence), consultable sur :

<http://www.ccsds.org/CCSDS/documents/650x0b1.pdf>

2.1.1. Réception des documents

Dans le cas des télétransmissions transitant par une plate-forme d'échanges, ces télétransmissions pourront parvenir aux services d'archives par réseau, parfois sous forme de flux XML.

La plate-forme d'archivage doit donc être notamment en mesure de recevoir ces flux, de vérifier leur provenance et leur intégrité, d'analyser leur conformité à des règles préétablies, de renvoyer à l'émetteur des notifications de réception ou d'anomalie en mode synchrone.

Certains documents pourront bien entendu être transmis selon d'autres modes : par messagerie ou par des supports amovibles (CD-R, bandes...) par exemple. Seront choisis les supports les plus couramment utilisés et répandus, afin de limiter les types d'appareils de lecture par la plate-forme d'archivage.

Une ou plusieurs stations de travail devront permettre à des opérateurs, aidés de logiciels, de vérifier les documents reçus (présence des métadonnées, lisibilité des données, etc.) et d'effectuer divers traitements (conversion de format par exemple, assignation d'un identifiant unique, empreintes à effectuer pour garantir l'intégrité des données...). La plus grande automatisation de ces tâches devra être recherchée.

Les métadonnées, généralement en XML (suivant la DTD EAD), devront pouvoir être prises en compte par le système d'information existant du service d'archives. Il pourra également arriver que l'administration versera des métadonnées qui ne seront pas au format EAD (schémas propres au processus métier, voire métadonnées non normalisées) : ce sera alors à la plate-forme d'archivage de réaliser la conversion des métadonnées au format EAD.

Il est attendu un haut niveau de sécurité afin que la plate-forme d'archivage dans son ensemble soit à l'abri des attaques extérieures, par la mise en place d'une zone démilitarisée (DMZ) par exemple.

Il est également attendu une haute disponibilité du système de réception, de sorte que les flux puissent être reçus 7 jours sur 7, 12 heures par jour. Une indisponibilité maximale d'une journée (pannes, maintenance) sera tolérée.

2.1.2. Stockage des documents

Les fichiers de données et les métadonnées sont transmis par les opérateurs en charge de la collecte, au service de stockage au sein duquel ils seront organisés dans une arborescence virtuelle.

Un interface devra par conséquent être mise en place avec le service en charge du stockage des documents, avec notamment la précision de la classe de service à laquelle appartient chaque ensemble de documents transféré : en matière de niveau de sécurité, de délais de restitution, des droits d'accès tenant compte à la fois de la nature du document (délai de communicabilité) et de la nature du demandeur (service versant, citoyen).

L'infrastructure de stockage variera évidemment en fonction du volume de données prévu. La solution retenue permettra une montée en charge progressive des volumes.

Une attention particulière sera portée aux supports de type magnétique ou magnéto-optique et aux architectures SAN.

Elle devra assurer un très haut niveau de sécurité physique des données (réplications et sauvegardes à prévoir pour éviter les pertes) et une excellente traçabilité des opérations effectuées.

Un lien doit être prévu et mis à jour entre les métadonnées descriptives présentes dans le système d'information du service d'archives et l'adresse physique du document dans le service de stockage. L'adresse physique du document dans le service de stockage doit être gérée par ce service. Chaque objet d'information (données et métadonnées) transmis au service de stockage doit être repéré par un identifiant unique défini au moment de la réception des documents. Le lien entre cet identifiant et l'adresse de stockage est également géré par le service de stockage.

Le coût des éventuelles migrations de supports à prévoir doit également être envisagé ainsi que de leur recyclage systématique au-delà d'un certain délai. Pour les gros volumes de données, elles devront pouvoir être automatisées avec la prise en compte d'une surveillance des supports (et des évolutions technologiques) : intégration d'éléments statistiques sur les supports permettant de suivre leur état, système de lecture aléatoire par échantillonnage pour permettre la surveillance des supports. Ces migrations de support peuvent, suivant les technologies, constituer des activités ponctuelles périodiques et des activités continues intégrées au système lui-même.

Une hiérarchisation des supports de stockage peut être prévue, avec passage automatique d'un type de support à un autre, en fonction de la fréquence de consultation des documents et du recyclage nécessaire des médias.

2.1.3. Restitution des documents

La restitution des documents peut être à destination du service qui les aura confiés au service d'archives (soit directement, soit par l'intermédiaire d'une plate-forme d'échanges) ou à destination d'un citoyen intéressé.

Une interface devra par conséquent exister avec le service en charge du stockage.

La recherche du document souhaité sera réalisée par l'intermédiaire du système d'information du service d'archives.

La restitution peut s'effectuer par réseau (comme pour la réception des flux XML), par messagerie ou par envoi de supports amovibles. On devra par conséquent prévoir des capacités de copie d'objets numériques sur des médias de diffusion.

L'étude devra envisager et évaluer ces trois modes de restitution.

Un haut niveau de sécurité devra évidemment être assuré pour éviter toute intrusion dans la plate-forme d'archivage.

Une gestion fine des droits d'accès devra être mise en œuvre, prenant en compte les caractéristiques de communicabilité des documents saisis lors du stockage des documents.

Le délai de restitution acceptable dépend de la nature des documents archivés mais n'est généralement pas un élément déterminant.

Une partie des documents devra rester accessible en ligne immédiatement en permanence (documents numérisés très consultés par exemple).

Pour les autres documents, un délai de restitution de 15 minutes à quelques heures, voire même de plusieurs jours (selon l'effet sur les coûts), est envisageable.

Par ailleurs, des interruptions de service accidentelles peuvent être admises jusqu'à une journée.

2.2. Types de services d'archives et hypothèses de volumes de données

Il est difficile d'évaluer les volumes qui seront archivés dans la mesure où ce volume sera certainement très réduit au départ (et ne représentera qu'une partie minime par rapport aux images issues des fonds numérisés) et ne croîtra de façon exponentielle que progressivement (dans les 10 à 15 ans). Devra par conséquent être privilégiée une approche progressive : architecture de stockage permettant d'augmenter de la manière la plus souple possible la volumétrie des données archivées.

Le volume de données variera aussi en fonction de la taille du service d'archives.

Peuvent être distinguées quatre types de services d'archives :

- les petits services
- les services moyens non chefs-lieux de région
- les services chefs-lieux de région ou les gros départements non chefs-lieux de régions
- les gros services : Archives nationales, Archives départementales de Paris, du Rhône, des Bouches-du-Rhône, du Nord....

La moyenne des entrées d'archives annuelles sur support papier (en mètre linéaire) peut être répartie en quatre catégories : moins de 100 ml ; 100 à 350 ml ; 350 à 500 ml ; plus de 500 ml (3 à 4 services peuvent dépasser 1 km par an).

De la même façon, on peut imaginer plusieurs paliers en terme de volumes en octets pour les seules archives électroniques (sans compter les archives numérisées). Pour une première phase :

- moins de 5 TO
- de 5 à 50 TO
- plus de 50 TO

Ces seuils pourront naturellement être ajustés en cours d'étude.

Par ailleurs, on considèrera que la taille des objets archivés sera en moyenne autour de 1 à 5 MO.

2.3. Coûts à évaluer

Pour chaque type de service d'archives, chaque hypothèse de volume de données et chaque hypothèse de fonctionnalités à mettre en œuvre, l'ensemble des coûts de mise en place (développements) et de maintenance et fonctionnement de la plate-forme d'archivage devront être évalués, qu'il s'agisse de coûts humains, de matériel, de logiciels, de prestations de service...

Les différents postes de coûts seront soigneusement distingués entre :

- Les outils matériels et logiciels qui relèvent de la collecte et de la préparation des données

- Les outils matériels et logiciels qui relèvent du stockage

- Les outils matériels et logiciels qui relèvent de la restitution et communication.

De même, les coûts humains (nombre et profils des personnes : par catégories d'activités, niveau de compétence requis) et les coûts financiers seront distingués.

Les services d'archives possédant déjà parfois des infrastructures réseaux d'accès à des ressources électroniques (inventaires et fonds numérisés), ainsi que des outils de gestion et de description de leurs fonds d'archives papier, dans lesquels les données électroniques archivées pourraient prendre place, il devra être procédé, en outre, à une évaluation du simple coût de mise à niveau de ces services d'archives pour aboutir à une plate-forme d'archivage conforme aux caractéristiques définies plus haut.

Une évaluation sera donnée suivant :

- qu'un service n'a actuellement aucune infrastructure réseau et ne possède aucun outil logiciel de gestion et description de ses archives

- qu'un service possède une infrastructure réseau permettant d'accéder à des ressources électroniques, ainsi que des applicatifs permettant la publication et la recherche de ces ressources

- qu'un service possède également une infrastructure de stockage adaptée.

Afin de procéder à ses estimations, le prestataire retenu devra visiter un certain nombre de services d'archives territoriales (archives départementales, grands services d'archives municipales, archives régionales), correspondant aux types définis plus haut, ainsi que le service Constance du Centre des archives contemporaines de Fontainebleau. La Caisse des dépôts et consignations pourra également être contactée (jerome.bordier@caissedesdepots.fr) ainsi que le Conseil général des Yvelines (service informatique) (Clambert@cg78.fr). Il est également demandé de prendre contact avec les services du CNES Toulouse (service valorisation des données et outils d'ingénierie (claud.huc@cnes.fr)).

Une variante sera également estimée : le coût de mise en œuvre d'un centre informatique central, pour la partie *stockage*, cette fonction intégrant bien évidemment les migrations nécessaires qu'elles affectent les supports ou les formats des données, soit en direct par les Archives nationales, soit en sous-traitance, permettant la mutualisation de la fonction stockage, chaque service d'archives conservant la maîtrise des fonctionnalités liées à la collecte et à la restitution des données archivées. Le coût de ce service sera estimé en fonction du volume des données archivées.

2.4. Délai de réalisation

Le délai de réalisation de l'étude est de quatre mois à compter de la date de la commande.