Section 11 - Gestion d'un projet d'archivage électronique

Françoise BANAT-BERGER

Claude HUC

version 1

22 novembre 2011

Table des matières

I Section 11 - Gestion d'un projet d'archivage électronique

1. Chapitre 1. Objet de la section

Cette section se rapporte à la gestion d'un projet d'archivage électronique qui, comme tous les projets liés à la mise en œuvre d'un système d'information, doit s'accompagner d'une réflexion sur la gestion des risques et la maîtrise des coûts, avant le lancement à proprement parler du projet.

Par ailleurs tout projet d'archivage électronique doit s'accompagner de l'élaboration d'une politique d'archivage, spécifiant les rôles et responsabilités des différents acteurs et s'articulant elle-même, dans une politique de sécurité à mettre en œuvre, qui s'intègre dans la politique générale de sécurité de l 'organisme.

2. Chapitre 2. Gestion des risques

La méthodologie de gestion de risques est ancienne dans le monde industriel et « à la mode » dans le monde du management et des services.

Or, ***même si la préservation des documents numériques implique un ensemble d'actions préventives, il s'agit avant tout et un projet comme un autre.***

Que penser d'un système judiciaire qui n'aurait pas su conserver tous les documents nécessaires au bon déroulement d'un procès ?

Que penser d'un système de santé qui aurait perdu la trace des analyses, examens, radiographies antérieurs d'un patient ?

Si des séquelles d'un accident de la route apparaissent dix ou vingt ans après cet accident, que penser d'une société d'assurance qui n'aurait pas su pérenniser votre dossier ?

Que dire d'une grande agence scientifique qui conduit des projets sur la base de financements publics et qui n'aurait pas su conserver les données qui résultent de ces projets ?

Quelle que soit l'organisation concernée, la prise en compte du problème de pérennisation des informations numériques est d***e la responsabilité des dirigeants de cette organisation***. La gestion des risques est essentielle car elle fournit un ***référentiel*** sur lequel peut reposer l'élaboration des ***conventions de service entre producteurs**** ***et service d'archives**** ***pour le transfert de responsabilités*** :

- Elle fait en sorte que l'engagement de l'Archive en termes de conservation et d'accès soit proportionné aux les risques qui pèsent sur les objets,
- elle définit les conditions d'une mission continue de veille et de surveillance des risques, définie dans le modèle OAIS comme «***planification de la pérennisation*** »

⚠ *Attention*

La non prise en compte du problème conduit à terme à la perte d'informations qui peuvent être précieuses, voire indispensables au fonctionnement de l'organisation. Au mieux, la reconstitution d'une information perdue sera nécessaire et nécessitera des ressources considérables.

Ce terme pourra se situer entre cinq et quinze ans suivant les cas.

2.1 – Méthodologie et principes

La première étape consiste à identifier et classifier les risques. Identifier et classifier les risques, c'est posséder un référentiel qui les catégorise.

Pour ce faire, une bonne pratique est de constituer un ou plusieurs groupe(s) de travail réunissant les « spécialistes » ou du moins les personnes ayant la compétence métier concernée. Ce ou ces groupes auront en charge de dresser la liste des risques. Il s'agit alors d'analyser les aspects fonctionnels, mais aussi de disposer des points de vue technique ou administratif. L'aide d'une personne en charge de la qualité – quand cela est possible – constituera une aide précieuse pour la rigueur et le bon suivi de la démarche.

Une fois la liste des risques constituée, il faut les évaluer. Pour cette étape, il est possible d'avoir une approche AMDEC (Analyse des Modes de Défaillance, de leurs Effets et leurs Criticité) du problème.

À partir de l'évaluation des risques, il est possible d'identifier les risques prioritaires, notamment en fixant un seuil pour l'Indice de Priorité des Risques (IPR), seuil au-delà duquel un effort de maîtrise doit être entrepris.

2.2 – Méthodologie et principes : deuxième étape

Deuxième étape après la définition du risque :

L'évaluation du risque : selon son degré de probabilité (5 niveaux de « improbable » à « fréquent »), son impact (toujours 5 niveaux de « insignifiant » à « catastrophique ») et son degré d'apparition dans le temps (de « lointain » à « imminent »).

La prise de décision : les risques sont-ils acceptables ? Suivant les objectifs de l'organisme, selon le type de producteurs et de catégories d'utilisateurs, suivant les coûts et bénéfices des opérations de maîtrise des risques, suivant les obligations légales.

Les maîtrises sont les mesures à prendre pour endiguer le risque :

- qui diminuent la probabilité d'un risque (choisir des formats ouverts par exemple),
- qui diminuent l'impact d'un risque (avoir un plan de sauvegarde par exemple),
- qui évite le risque (on ne prend pas en charge des objets d'archives dont la conservation s'avère trop risquée).

Deux autres actions sont possibles :

- partage du risque : ainsi, dans le secteur privé, les tiers archiveurs souscrivent des polices d'assurance contre les atteintes aux informations et/ou les dommages immatériels, comme la destruction volontaire ou involontaire des données, la divulgation d'informations, la mauvaise qualité des programmes,
- ou encore la tolérance au risque : on décide par exemple d'encourir le risque et de payer des dédommagements plutôt qu'assurer la préservation sur le long terme de certaines des ressources numériques qu'on conserve.

Une itération s'avère ensuite indispensable pour vérifier que les maîtrises mises en œuvre n'ont pas créé d'autres risques, ou encore pour adapter l'évaluation des risques lorsque le système est modifié ou pour prendre en compte les risques évolutifs.

Complément : la démarche AMDE

AMDE propose une démarche de recherche déductive et exhaustive des risques encourus par un système. Ici, il s'agit d'une recherche des causes en mesure d'entraîner une défaillance, d'une recherche des dispositions existantes en mesure de détecter la cause, d'une recherche des recommandations permettant de réduire voire de supprimer la cause ou son impact. En ajoutant la criticité, AMDEC permet une hiérarchisation des risques.

Plusieurs moyens peuvent être utilisés pour définir le niveau de criticité. On en retiendra deux qui permettent de fixer un Indice de Priorité des Risques (IPR) au-dessus duquel toute criticité doit être réduite :

- la multiplication d'un indice de gravité par un indice d'occurrence de la cause,

- la multiplication d'un indice de gravité par un indice d'occurrence de la cause et par un indice de détection des contrôles.

Par exemple, nous pouvons évaluer la gravité d'un risque graduellement depuis le niveau « sans gravité » à « catastrophique » sur une échelle de 0 à 5, l'occurrence de la cause évoluant graduellement de « peu probable » à « très fréquent » de 1 à 5 et l'indice de détection des contrôles de « facilement détectable » à « très difficile » à détecter de 1 à 5.

Prenons l'exemple du risque de dégradation des supports dans le cas d'un enregistrement effectué sur un support de type CD-R dans le respect des préconisations en vigueur. L'indice de gravité est estimé à « grave » et non « catastrophique » (4) car il existe également une autre copie de sauvegarde, l'indice d'occurrence est estimé à « moyennement fréquent » (3) car la durée d'un CD-R est de l'ordre de 3 à 5 ans et l'indice de détection des contrôles est estimé au « maximum » (5) car aucun moyen de contrôle n'a été mis en place. Ainsi nous obtenons une évaluation du risque de 60 (IPR=60). Nous avons les moyens d'identifier qu'il est donc primordial de détecter la dégradation des CD-R.

Complément : la méthode « Ishikawa »

Cette méthode est dite de l'arête de poisson due à la forme du diagramme obtenu – issue du domaine de la gestion de la qualité. Le but est de trouver toutes les causes possibles d'un effet ou d'un problème. Cette méthode graphique permet de faciliter la réflexion (« brainstorming »).

Chaque arête du diagramme constitue une dimension du problème qu'il convient d'explorer. La version initiale de Kaoru Ishikawa propose cinq axes, « les 5M » : Matière, Matériel, Méthode, Milieu et Main-d'œuvre mais d'autres déclinaisons ont été adaptées à des secteurs d'activités particuliers. Nous proposons sur la figure 6.1 un exemple dans une version « 7M » d'un problème crucial : « la perte d'un fichier sur un support »

Perte d'un fichier sur bande

✎ *Remarque*

L'entité de management (au sens du modèle OAIS)

- c'est elle qui définit le mandat

- c'est en général elle qui fournit les ressources.

Ce n'est pas pour autant elle qui assure la gestion ou l'organisation de l'Archive.

2.3 – Classification des risques

Les risques environnementaux sont les risques globaux encourus par l'environnement de l'Archive, à savoir :

- les risques naturels,
- la sécurité,
- les lieux de stockage.

L'environnement est défini comme l'interaction des éléments matériels au sein desquels le système est installé :

- le lieu géographique, (il s'agit des risques naturels liés à un événement naturel),
- le bâtiment. (des risques de sécurité (risques liés à l'incendie ou à l'intrusion) et enfin les risques liés aux locaux.

Les risques organisationnels sont parfois largement sous-évalués voire ignorés. Pour ce faire, il est nécessaire de procéder à une ***réflexion sur les compétences nécessaires***. Les risques budgétaires constituent également un écueil important : si les coûts liés à la mise en place technique de la plateforme sont généralement prévus, en revanche les coûts de maintenance et de fonctionnement de la plateforme ne sont pas toujours intégrés dans le budget du service. Et surtout, sont généralement ignorés les coûts liés à la prise en charge d'un nouveau processus auprès d'un producteur. Cette catégorie de risques recouvre également les risques liés aux personnels (manque de compétences techniques par exemple), ainsi que les risques liés aux processus (manque de surveillance, mauvaise application des procédures de tests…).

Les risques technologiques concernent

- d'une part les supports d'enregistrement (obsolescence technologique et dégradation des supports),
- d'autre part les formats de représentation (obsolescence technologique du format de représentation lui-même).

Les risques liés à l'accès :

- risques liés à l'accessibilité sémantique (risques liés à la compréhension de l'Objet information auquel on donne accès à une communauté d'utilisateurs, ainsi que les informations nécessaires pour trouver, reconnaître et identifier les objets,
- risques liés à l'accessibilité technique (risques liés à la capacité technique de diffuser une information à une communauté d'utilisateurs : absence de métadonnées techniques ou de structure appropriées, présence de systèmes de protection, cryptage des données),
- risques liés aux responsabilités d'accès de l'Archive (risques liés à l'organisation, aux informations et aux outils nécessaires pour qu'un objet Contenu d'information ne soit accessible qu'aux personnes qui disposent des droits d'accès à cet objet).

⊕ *Complément : la nécessaire implication des dirigeants de l'organisme et les compétences nécessaires au projet*

L'archivage numérique est une activité à part entière à la croisée entre archivistique et informatique.

Le succès de ce type de projet réside dans la capacité de ses acteurs à travailler ensemble ou dans l'émergence de nouveaux profils rassemblant ces deux compétences.

Étant donné la diversité des compétences sollicitées dans la construction du projet (archivistiques, informatiques, juridiques, qualité) et, généralement, l'absence de structure ou service spécialisé dans ce domaine, ***la nomination d'un groupe de travail ou la création explicite d'un projet par la direction est un signe qui permet de rendre visible à tous l'importance de ce projet***. Cette solution permet de montrer la nature transverse du problème aux acteurs internes, qui ont forcément une approche du point de vue de la problématique de leur métier, et aux acteurs externes qui vont apporter soit une formation pour pallier le manque de compétence partielle des équipes, soit une expertise lorsqu'il s'agit de compétences qu'il n'est pas nécessaire d'acquérir sur le long terme.

De plus, la mise en place d'un projet d'archivage électronique implique de faire évoluer les processus existants ou d'en mettre en place de nouveaux, notamment pour collecter les métadonnées nécessaires à la pérennisation. Pour obtenir l'adhésion des services et des personnes concernés par ces changements, il est nécessaire de la part de l'organisation d'afficher une volonté sans laquelle rien ne sera vraiment possible.

Pour aller plus loin, il faut envisager la création d'équipes de projets et de structures permanentes entièrement dédiés à l'archivage numérique, et faire en sorte que ces équipes disposent des différentes compétences requises, et enfin les organiser de façon à ce que ces différentes compétences reposent toutes sur une même compréhension du problème à résoudre.

Nous le voyons, le rôle de la direction est déterminant. La réussite du projet repose en partie sur l'affichage qu'elle montre à soutenir le projet. C'est pourquoi un tel projet doit avoir des objectifs clairs et visibles par l'ensemble des intervenants.

2.4 – Conclusions sur la gestion des risques

Il s'agit d'utiliser les données de gestion des risques pour générer un ***tableau de bord de pilotage de la préservation des objets numériques d'un secteur (« plan de conservation** »).

C'est un outil de pilotage qui permet de :

- planifier les ***revues*** : surveiller l'évolution des données, des applications et des supports. La revue permet de s' assurer que les choix techniques ou organisationnels sont toujours valides et que leur qualité est bonne. La revue produit un compte-rendu de revue qui déclenche une alerte en cas de situation anormale,
- prévoir les alertes et les solutions : un signal est émis suite à la détection d'un problème, l'alerte déclenche un processus de résolution des problèmes préparé à l'avance (plan d'urgence…),
- déterminer les risques spécifiques à différents types d'objets, et pondérer l'évaluation des risques en fonction de leurs particularités techniques ou autres.

On peut également auditer le système en vue de la certification, ou en vue de prioriser les actions. La gestion des risques donne une vue d'ensemble des actions de préservation menées sur l'ensemble des objets numériques.

3. Chapitre 3 - Maîtrise des coûts

Tout projet de mise en œuvre d'une Archive numérique* doit naturellement répondre aux exigences méthodologiques de tout projet en matière d'identification des étapes et des conditions requises pour passer d'une étape à la suivante. Un certain nombre d'éléments préalables doivent être établis avant de rentrer dans le détail de l'analyse.

- L'entité de Management de l'Archive est clairement identifiée. Elle est l'autorité qui définit le mandat de l' Archive, et lui affecte les ressources nécessaires à l'accomplissement des objectifs découlant du mandat,
- Le mandat donné à l'Archive par le Management est clairement explicité : qui sont les producteurs, qui sont les utilisateurs, quels sont les services qui seront offerts à ces derniers ? quels sont les objets à archiver, leur durée de conservation, leur force probante ?
- Les différents aspects réglementaires et juridiques ont été explicités par écrit,
- L'Archive dispose de ressources quantifiées,
- L'Archive dispose d'une politique d'archivage,
- L'Archive doit élaborer un plan de réversibilité, au cas où l'Archive devra mettre fin à ses activités (possibilité d' extraire l'ensemble des données et métadonnées vers une autre Archive).

3.1 – L'évaluation des coûts

L'évaluation de ces coûts est et restera un exercice difficile et cependant indispensable. Le développement croissant des technologies numériques a pour conséquence une croissance présente et future des volumes de données qu'il convient d'archiver. Cette croissance est bien plus rapide que n'a pu l'être celle des archives papier.

L'entité responsable de l'Archive ou son donneur d'ordre ne saurait s'engager dans un cycle de coûts continûment croissants sans espoir de stabilité et encore moins de réduction de ces coûts. Il devient nécessaire non seulement d'être capable d'évaluer ce que seront les coûts de l'Archive, mais aussi de montrer que l'Archive sera capable (sauf en cas de modification significative de son mandat), de fonctionner avec des coûts constants dans un contexte d'augmentation constante des volumes d'information à archiver.

Exemple

La NASA, riche de son expérience en matière de traitement et d'archivage de données, tente de mettre au point un modèle d'évaluation des coûts de développement, d'exploitation et de maintenance des systèmes de traitement et d'archivage de données scientifiques. Un logiciel libre, le « Cost Estimation Toolkit » est distribué et ne demande qu'à être expérimenté.

Lien vers le site de la NASA : « Cost Estimation Toolkit and Comparables Database »

http://opensource.gsfc.nasa.gov/projects/CET/CET.php

Complément : le projet britannique LIFE

Le projet britannique LIFE (Life Cycle Information for E-Literature, *http://www.life.ac.uk/*) propose une intéressante modélisation du cycle de vie de l'information numérique visant à calculer les coûts de la conservation de cette information sur une durée allant de cinq à dix ans. Cette modélisation, orientée vers les bibliothèques numériques, a été validée sur trois exemples réels qui sont :

- une collection de 170 objets provenant du dépôt numérique volontaire depuis 2001 à la British Library,
- les activités d'archivage du Web portant sur 1000 sites par an dans le cadre de la contribution de la British Library au consortium UKWARC (United Kingdom Web Archiving Consortium),
- les journaux électroniques à la bibliothèque de l'UCL (University College London) qui gère des abonnements portant sur 12 365 revues périodiques.
  LIFE propose l'approche analytique suivante : le coût de l'archivage peut être évalué en analysant de façon distincte les coûts qui ne sont pas liés au cycle de vie de l'information (management, administration, infrastructure) et les coûts directement liés au cycle de vie des objets numériques. C'est sur cette partie que l'analyse est la plus pertinente avec une décomposition du coût C comme suit :
  C = Aq + I + M + Ac + S + P
  Dans cette équation :
  Aq (acquisition) coût correspondant à la collecte et récupération des objets,
  I (Ingest) coût correspondant à l'évaluation, l'analyse et la validation des objets et à leur insertion dans le dépôt numérique,
  M (Metadata) coût correspondant à la création, à l'extraction et à l'enregistrement des métadonnées,
  Ac (Access) coût correspondant aux fonctions d'accès aux objets pour la communauté des utilisateurs,
  St (Storage) coût correspondant au stockage et au maintien de l'intégrité physique des objets,
  P (Preservation) coût correspondant aux activités de la fonction Planification de la pérennisation du Modèle OAIS.

3.2 – L'étude conduite pour les services publics d'archives

Dans une étude des coûts, il est intéressant d'analyser en particulier ***l'impact du modèle d'organisation retenu*** et par conséquent ***l'impact de la mutualisation des moyens sur les coûts***. C'est ce que peuvent nous démontrer des études conduites, par exemple, pour les services publics d'archives.

Exemple : Étude conduite en France

Une étude relative à la réalisation de plates-formes d'archivage électronique pour les services publics d'archives, conduite par la société Parker Williborg à la demande de la Direction des Archives de France, aborde la question de l'évaluation des coûts dans la sphère publique avec ***un niveau de détail intéressant***.

Elle analyse en particulier l'impact du modèle d'organisation retenu et par conséquent ***l'impact de la mutualisation des moyens sur les coûts***.

Cinq scénarios ont été étudiés, depuis une plate-forme locale et isolée jusqu'à une plate-forme nationale et largement mutualisée :

- une plate-forme dédiée à un seul service producteur qui assure lui-même son archivage ou dispose d'un service d'archive interne (exemple d'une grande municipalité, d'un conseil général) : scénario 1
- une plate-forme dédiée à un ensemble de services producteurs locaux (comme les services d'archives départementales ou certains gros services d'archives municipales) : scénario 2
- une plate-forme nationale dédiée à l'ensemble des services centraux et déconcentrés relevant d'un même ministère : scénario 3
- une plate-forme nationale dédiée à un ensemble de collectivités de même type (conseils généraux, conseils régionaux) : scénario 4
- enfin une plate-forme nationale pour l'ensemble des administrations centrales de l'Etat (Archives nationales) : scénario 5.

Exemple

L'estimation des volumétries attendues pour les 10 ans à venir est basée sur l'identification des sources candidates à l'archivage électronique. Elle tient également compte de la conservation de toutes les données en deux exemplaires.

L'estimation des ressources humaines nécessaires est basée sur une analyse des fonctionnalités de la plate-forme, par grands processus (préparation et prise en charge des versements, stockage, gestion des données descriptives, restitution) et par fonctions transverses (administration de la plate-forme, pilotage, veille technologique et juridique, projets d'évolutions et de migrations). Les tâches ont été ensuite décomposées avec la détermination d'un temps moyen par tâche et un coût horaire des agents de catégories A (bac + 3) et B (bac à bac +2). Des estimations différentes ont été dressées suivant le choix d'une plus ou moins grande automatisation des processus, l'utilisation, pour les transmissions, de réseaux ou de supports amovibles, la mise en place de plates-formes manuelles (supports sur rayonnages) ou automatisées (juke-box, librairies, baies de disques…).

L'évaluation des coûts initiaux et des coûts annuels d'exploitation de la plate-forme d'archivage électronique a été basée sur l'hypothèse suivante, à savoir le développement au niveau national, sur la base de progiciels du marché (pour lesquels un travail d'intégration important sera demandé), d'une solution générique en passant par les étapes suivantes :

- développement,
- réalisation d'un déploiement pilote,
- généralisation de la solution.

C'est ce qui a conduit au ***développement de la plate-forme pilote PIL@E***.

La plate-forme générique étant supposée disponible, les coûts spécifiques à chaque plate-forme déployée ont alors également été estimés. Ces coûts comprennent d'une part les ***coûts initiaux*** (acquisition de la plate-forme d'exploitation, installation de la solution générique adaptée, coûts internes de démarrage auprès des services producteurs) et d'autre part ***les coûts d'exploitation annuels*** suivant les processus (prise en charge des versements, coût de gestion du stockage, coût des restitutions et consultations, coût de maintenance applicative et technique, coût des fonctions transverses).

Enfin, ***des synthèses*** ont été réalisées scénario par scénario ainsi que ***des comparatifs des coûts*** au téraoctet.

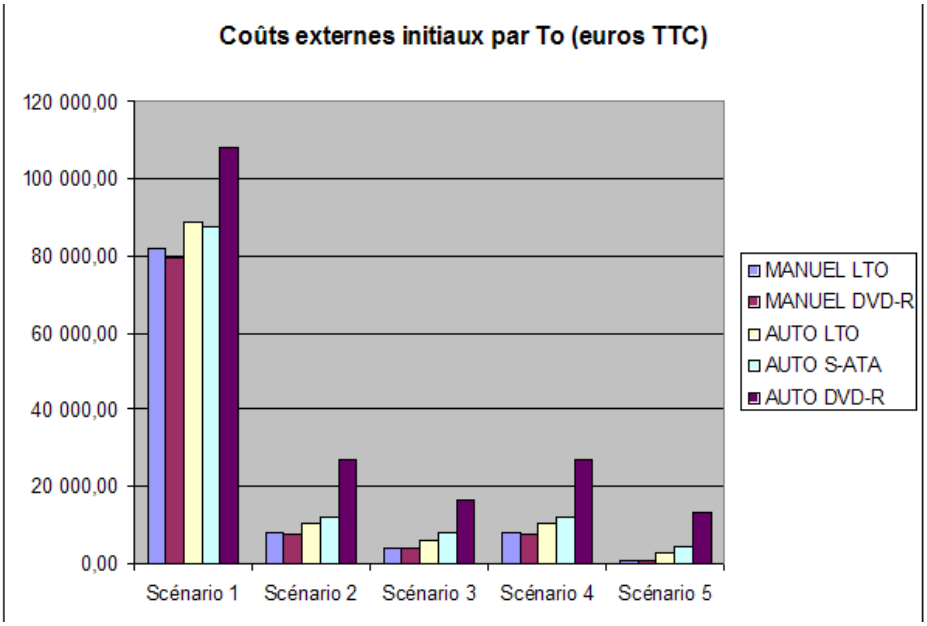
Il en ressort que dès lors que les volumes archivés augmentent, les économies d'échelles sont considérables :

• Les ratios entre les coûts par téraoctet sont systématiquement supérieurs à vingt entre le premier scénario (plate-forme d'archivage dédiée à un seul service producteur) et le dernier (plate-forme nationale) que ce soit pour les coûts initiaux externes ou les coûts d'exploitation. ***Il est par conséquent préconisé d'encourager les plates-formes d'une certaine ampleur***.

• Les coûts internes les plus importants sont ceux afférents au démarrage d'un processus d'archivage avec un service producteur (nouveau producteur/nouvelle application) avec une forte charge sur les catégories A au démarrage et sur les catégories B pour l'exploitation. Ceci vaut tant pour le service d'archives que pour le producteur qui devra adapter son application pour le transfert au format défini par le schéma XML de versement vers la plate-forme d'archivage.

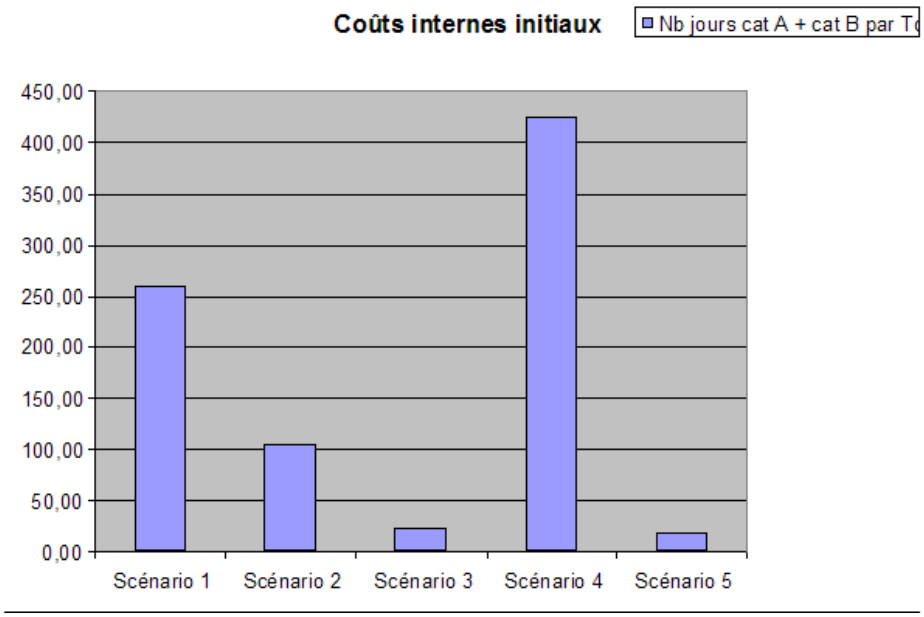
Par conséquent, il vaut mieux privilégier les versements de gros volumes et limiter autant que possible le nombre de services producteurs différents, si on ne dispose pas de ressources humaines suffisantes.

• De même, l'étude préconise fortement, dans un souci de mutualisation, de renforcer l'échelon central (Direction des Archives de France) afin d'apporter une aide à la modélisation des versements par grandes catégories de documents, que l'on retrouve sur tous les sites (exemple des marchés publics).



évaluation des coûts externes

Expliciter quelque part la signification des types de support LTO, S-ATA n'est pas clair pour tout le monde



évaluation des coûts initiaux en Euros.

Titre du tableau : Cas d'un seul service producteur qui assure lui-même son archivage

	PLATE-FORME MANUELLE / LTO		
	Coûts externes (euros TTC)	Agent catégorie A (j/h)	Agent catégorie B (j/h)
Coûts initiaux			
Acquisition Plate-forme Exploitation	39 468,00		
Installation et déploiement sur site	42 338,40	50,00	
Démarrage des services producteurs		10,00	
Démarrage des applications sources		200,00	
Total	**81 806,40**	**260,00**	**0,00**
Coût au To	**81 806,40**	**260,00**	**0,00**
Coûts d'exploitation annuels			
Versements			27,58
Stockage	388,70		0,23
Consultations/Restitutions	299,00		10,15
Maintenance applicative	8 467,68	10,00	
Exploitation et maintenance technique	10 405,20		40,00
Fonctions Transverses*		35,00	95,00
Charges variables de pilotage**		21,80	
Total	**19 560,58**	**66,80**	**172,96**
Coût au To	**19 560,58**	**66,80**	**172,96**

* Hors charges variables de pilotage
** égal à 10% des effectifs (agents catégories A et B)

Cas d'un seul service producteur qui assure lui-même son archivage

	PLATE-FORME AUTOMATISEE / LTO		
	Coûts externes (euros TTC)	Agent catégorie A (j/h)	Agent catégorie B (j/h)
Coûts initiaux			
Acquisition Plate-forme Exploitation	222 755,00		
Installation et déploiement sur site	42 338,40	50,00	
Démarrage des services producteurs		200,00	
Démarrage des applications sources		1 600,00	
Total	**265 093,40**	**1 850,00**	**0,00**
Coût au To	**2 650,93**	**18,50**	**0,00**
Coûts d'exploitation annuels			
Versements			220,60
Stockage	38 870,00		
Consultations/Restitutions	5 980,00		78,00
Maintenance applicative	8 467,68	10,00	
Exploitation et maintenance technique	47 062,60		40,00
Fonctions Transverses*		35,00	190,00
Charges variables de pilotage**		57,36	
Total	**100 380,28**	**102,36**	**528,60**
Coût au To	**1 003,80**	**1,02**	**5,29**

* Hors charges variables de pilotage
** égal à 10% des effectifs (agents catégories A et B)

Cas d'une plate-forme nationale pour l'ensemble des administrations centrales de l'État

3.3 – Les facteurs à prendre en compte dans l'évaluation des coûts

Eléments à prendre en compte	Ressources humaines		infrastructure - locaux, machines, réseaux,		développement et validation logiciels	
	Mise en œuvre	Exploitation	Mise en œuvre	Exploitation	Mise en œuvre	exploitation
Nombre de services producteurs (nombre de protocoles de versement, nombre d'interlocuteurs)	Fort	Fort				
Nature des relations avec les producteurs : est-il possible d'imposer les formats d'archivage ou est-ce que l'Archive sera contrainte à des migrations de format dès le départ, fourniture complète ou non de métadonnées	Fort				Fort	
Complexité des contenus (nécessitant d'associer des éléments documentaires aux « objets information » archivés)	Fort	Fort				
Nombre et complexité des formats de données	Fort				Fort	Fort
Proportion des documents soumis à des contraintes de valeur probante (impliquant un usage systématique de signatures électroniques, de calculs d'empreintes et d'horodatage)	Fort		Fort		Fort	
Faible niveau d'automatisation des versements (existence éventuelle d'opérations manuelles systématiques)		Fort				
Haut niveau d'automatisation des versements					Fort	Fort
Disponibilité de composants logiciels réutilisables					Fort	
Volumétrie élevée (impact sur la capacité des moyens de stockage, sur les opérations de surveillance et leur renouvellement régulier dans le cadre des migrations de support)		Moyen	Fort	Fort		
Granularité des objets (la gestion d'un très grand nombre d'objets n'est pas sans impact sur les contraintes pesant sur la base de données)			Fort		Moyen	
Exigences relatives à la sécurité, à la criticité et la confidentialité des données	Fort		Fort		Fort	
Exigences sur la continuité du service : taux d'indisponibilité acceptables. Un service ouvert 24h/24h avec un taux d'indisponibilité faible doit disposer de moyens redondants et d'un système d'astreinte		Fort	Fort	Fort	Moyen	Moyen
Nombre d'utilisateurs pouvant consulter le service en parallèle			Fort			
Exigences sur le niveau de service aux utilisateurs : accéder aux documents plus ou moins rapidement, interface et moyen de recherche et de récupération sophistiqués			Fort		Fort	
Service client (gestion des utilisateurs et de leurs droits d'accès)		Fort				Moyen à fort
Fréquence des audits sur les paquets versés et les paquets archivés		Fort				
Support gratuit aux utilisateurs (si le support est facturé, cela constituera une source de ressources pour l'Archive)		Fort				
Veille technologique, suivi des évolutions des standards, des besoins des utilisateurs...		Fort				

tableau synthétique des facteurs de coûts à prendre en compte

3.4 – Les pistes pour la réduction des coûts

Les premières pistes à examiner sont internes à l'Archive et pointent sur son organisation , son fonctionnement, ses équipements et la rationalisation de ses activités.

Ensuite, les possibilités de partager certaines dépenses avec d'autres Archives ou d'autres services de l'organisme sont fort diverses et nombreuses. L'objectif étant ici de répartir les coûts sur une base aussi large que possible :

- Réutilisation d'une infrastructure informatique existant dans l'organisme,
- Partage d'une infrastructure de stockage : l'expérience montre que mettre en place puis assurer le fonctionnement opérationnel et la maintenance d'une infrastructure de stockage va nécessiter les mêmes compétences et les mêmes ressources humaines, qu'il s'agisse de stocker 50 To ou 500 To.,
- Sachant qu'une Archive devrait veiller à ce que les objets numériques soient stockés sur deux sites distincts géographiquement éloignés, il pourra être opportun de rechercher des accords de réciprocité avec d'autres Archives plutôt que de gérer deux sites de stockage.
- Développement de systèmes logiciels génériques adaptables à de multiples contextes au sein d'un même domaine. C'est par exemple cette démarche qui est suivie dans un service d'archives qui reçoit des archives provenant de multiples producteurs mais qui utilise un seul outil de recherche pour l'ensemble des fonds archivés quels que soient les domaines administratifs,
- Réutilisation des composants logiciels de l'application d'archivage. Le développement des Archives numériques draine dans son sillage, des composants logiciels libres ou commerciaux qui doivent éviter à de nombreux sites d'archivage numérique de lancer et donc financer le développement de logiciels répondant à leurs besoins. On peut penser que dans un domaine donné, par exemple celui des services publics d'archives, celui des archives scientifiques ou d'autres, 90 % des besoins sont communs. Il s'ensuit que de nombreux logiciels devraient disposer d'un bon potentiel de réutilisation s'ils ont été conçus dans cet esprit, ce qui est d'ailleurs de l'intérêt des éditeurs de logiciels.

Dans le domaine du versement, deux points importants sont à considérer :

- le premier est la prise en compte des contraintes liées à l'archivage des documents, au moment de la création du document ou au moment où le document est figé et devient non modifiable.
- Le second est l'automatisation maximale du processus de versement (qui implique une standardisation des paquets SIP) et de création des Paquets d'information archivés.

Les travaux de définition de standards ne doivent en aucun cas être à la charge d'une Archive unique. Il conviendra de réutiliser les standards existants, et lorsqu'il y a des besoins nouveaux, de partager les tâches de rédaction de nouveaux standards avec d'autres Archives.

Exemple

C'est ainsi que la direction des Archives de France et la direction générale de la modernisation de l'État mettent à disposition de l'ensemble des services publics d'archives, non pas des composants logiciels mais des modèles de description conformes au standard d'échange de données pour l'archivage pour un catégorie d'archives qu'on retrouve sur tout le territoire, indispensables pour pouvoir développer les exports.

4. Chapitre 4. Elaboration d'une politique d'archivage

Avec le numérique, les risques de contentieux deviennent plus importants : détermination des rôles et responsabilités des différents acteurs intervenant tout au long du cycle de vie du document, problématiques d'intégrité (comment prouver que le document originel n'a pas été modifié ?), problématiques de conversion de formats (comment prouver que le document converti n'a pas perdu telle fonctionnalité d'origine), problématiques de supports défectueux qui ont endommagé ou fait disparaître la donnée, problématiques de droits d'accès.... C'est la raison pour laquelle il est

nécessaire de préciser dans le cadre d'une politique d'archivage, l'ensemble des éléments qui vont participer à la mise en œuvre de processus d'archivage propres à assurer qu'un document a été convenablement intégré, contrôlé, conservé, géré et consulté tout au long de son cycle de vie.

4.1. Objectifs de la politique ou charte d'archivage

L'action de l'archiviste sera d'autant plus efficace qu'une charte d'archivage existera au sein de l'organisation, validée au plus haut niveau et valable pour toute l'information reçue ou produite dans l'organisation, quelle que soit sa forme. La charte ou politique d'archivage devra par conséquent s'appliquer à l'ensemble de la production avec des spécificités bien évidemment liées à la production numérique.

Cette ***charte*** ou politique d'archivage énoncera les bonnes pratiques sur lesquels se fonde cet archivage. Elle précisera l'environnement juridique en vigueur, identifiera les acteurs en présence et détaillera leurs obligations et responsabilités respectives. Ainsi la politique d'archivage définit les exigences minimales, en termes juridiques, fonctionnels, opérationnels, techniques et de sécurité, qu'une Autorité d'archivage doit respecter afin que l'archivage électronique mis en place puisse être regardé comme fiable. Elle repose sur des contraintes « standard » à mettre en place. Contraintes en matière

- d'identification/authentification de l'origine des documents archivés,
- de l'intégrité des archives,
- de leur intelligibilité et lisibilité,
- de leur durée de conservation,
- de la traçabilité des différentes opérations (notamment versement, consultation, élimination)
- et enfin de la disponibilité et de l'accessibilité des archives.

Cette charte ou politique constitue par conséquent un référentiel de la sécurité de l'archivage électronique pour qu'il puisse être qualifié de « fiable ». Une grille d'audit constituée à partir de ses différents chapitres permet par ailleurs à un auditeur de contrôler la fiabilité d'un service d'archivage électronique. L'audit doit notamment vérifier que les processus de mise en application de la politique d'archivage ont été définis et que ces processus sont réellement appliqués. Le responsable de la mise en œuvre de cette politique d'archivage doit être désigné clairement et connu car les différents intervenants peuvent en avoir besoin à titre informationnel et opérationnel.

La charte d'archivage constitue un document de haut niveau qui a être vocation à être complété par des documents plus détaillés décrivant sa mise en application en termes d'organisation, de processus et de procédures.

Exemple

Voir sur le site de l'agence nationale de la sécurité des systèmes d'information (ANSSI), les référentiels élaborés en 2006 concernant la politique d'archivage sécurisé dans le secteur public : *http://www.ssi.gouv.fr/site_article48.html*

4.2 – Responsabilités et obligations des différentes parties

L'Autorité d'archivage est l'entité qui a la responsabilité de l'archivage (gestion, traitement, conservation et communication des données).

En fait tout au long du cycle de vie d'un document, il peut y avoir plusieurs Autorités d'archivage successives : dans le secteur public, le service producteur tant que le dossier est « vivant », puis un service d'archives intermédiaire pour l' « âge intermédiaire » et enfin un service d'archives définitif pour l' « âge définitif » ; ou bien le producteur durant les

âges vivant et intermédiaire, et le service d'archives pour l'âge définitif. Le transfert anticipé des données vers le service d'archives avant la fin de la période intermédiaire ne modifie pas cette répartition des responsabilités.

De son côté, le service informatique aura un rôle d'opérateur qu'il exercera dans un premier temps pour le service producteur et, dans un second temps, pour le service d'archives lorsque celui-ci sera devenu Autorité d'archivage.

Le contrat/convention ainsi passé entre les différents acteurs du processus (service producteur, service d'archives, service informatique) définit le périmètre des données à verser / à éliminer ainsi que les processus de versement /d'élimination.

L'analyse générale des interfaces et interactions entre le service versant et le service d'archivage électronique constitue l'objet de la norme ISO 20652 « Producer Archive interface methodology abstract standard ». Cette norme est étudiée dans la partie 5 sur le modèle OAIS et les normes dérivées.

Enfin, les autres acteurs du processus (contrôleurs, usagers) auront un certain nombre d'obligations et de responsabilités à définir dans le cadre de cette politique d'archivage, notamment en matière d'accès à l'information.

4.3 – Obligations du service producteur

Le service doit assurer :

- La gestion (alimentation et mise à jour) du système d'information, signalement au service informatique des doublons, erreurs, dossiers vides,
- L'intégration du cycle de vie des données numériques au système d'information,
- En cas d'existence d'un module d'archivage interne ou d'une base d'archivage intermédiaire, le transfert des dossiers clos et le signalement de cette base en cas de données à caractère personnel, ainsi que le maintien de l'intégrité des données dans cette base,
- Pour les transferts vers le service d'archives, les aspects de volumétrie et le cas échéant, le calendrier prévisionnel pour les transferts à venir,
- La définition des supports éventuellement employés dans le cas d'un transfert n'utilisant pas de réseau informatique et pour les transmissions par réseau, les protocoles à utiliser,
- La définition des formats d'encodage des données que le service d'archives accepte ou éventuellement les modalités de conversion de formats que le service d'archives souhaite réaliser à l'arrivée des archives transférées si leur format d'origine ne permet pas d'assurer une bonne pérennité de l'information.
- La fourniture de toutes les informations relatives à la nature et à la durée de vie des archives transférées ainsi que leur éventuel caractère confidentiel et/ou les accès limités. Ainsi, le service producteur est responsable de l'exactitude de ces informations et de leur bonne transmission,
- La conformité aux exigences techniques définies par l'Autorité d'Archivage (notamment, le respect d'un format d' échange de données, ou encore la garantie que les supports et les archives qu'ils contiennent sont en parfait état et exempts de tout virus ou autre dysfonctionnement susceptible d'avoir un impact sur la bonne exécution de la politique d'archivage),
- En cas de documents signés électroniquement devant être transférés au service d'archives, la vérification de leur signature avant le transfert prévu et les résultats de cette vérification portés dans les métadonnées des documents concernés,
- Lors du transfert, la génération d'empreintes des archives transférées afin de permettre à l'Autorité d'archivage de vérifier l'intégrité de cette dernière au moment de sa réception.

4.4 – Obligations du service informatique

En tant qu'opérateur d'archivage, il s'engage :

- à procéder concrètement pour le compte de l'Autorité d'archivage aux opérations de versement et/ou d'élimination des données numériques,
- à assurer la suppression physique des données numériques après expiration de la durée dite d'utilité administrative, et obtention du visa d'élimination par le service d'archives,
- à délivrer un procès-verbal de destruction pour les données éliminées,
- à assurer un accès sécurisé aux utilisateurs de l'Archive par la mise en place d'une gestion des habilitations appropriée et des moyens d'accès dimensionnés selon les besoins et les ressources,
- à assurer les développements des évolutions nécessaires,
- à procéder à l'administration et à l'exploitation des systèmes,
- à garantir la sûreté de fonctionnement des systèmes par la mise en place d'un plan de reprise et/ou de continuité d'activité selon le niveau de qualité de service requis par les utilisateurs,
- à assurer l'intégrité des objets numériques par un stockage sécurisé des données numériques (redondance, réplication sur plusieurs sites distants, surveillance des supports et migration de ces supports),
- à assurer la veille technologique d'évolution de l'infrastructure,
- à mettre en œuvre les opérations de migration demandées par le service d'archives (prototypage, test, suivi d'exécution, ...).

4.5 – Obligations du service d'archive

Le service d’archive :

- définit, en collaboration avec le service producteur, les règles relatives au cycle de vie des données dans le système d'information,
- instruit les demandes de visas d'élimination qui lui sont adressées,
- reçoit, dans le format d’échange spécifié, les archives transférées et les contrôle pour validation ou rejet. Si le contrôle s'avère satisfaisant, production d’un message d'acceptation du transfert pouvant éventuellement être revêtu d'une signature électronique, qui marque la prise en charge et par là-même la responsabilité de l'Autorité d'archivage sur les archives transférées,
- applique, en cas de transfert avant expiration de la durée d’utilité administrative (DUA), la durée de conservation adéquate pour les archives concernées conformément aux instructions données par le service producteur,
- définit les règles permettant une conservation pérenne des données, à mettre en œuvre par le service informatique,
- vérifie l’ensemble des accès au service d’archives tant physiques que logiques, internes et externes en fonction des droits de chacun des intervenants, et à ne permettre les accès aux archives traitées qu’aux seules personnes habilitées,
- demande au service producteur, l'autorisation de celui-ci pour toute demande de communication des archives concernées, dès lors que les délais de libre communicabilité ne sont pas expirés,
- prévoit la signature d’un accord de secret et de confidentialité par tout personnel externe et par les éventuels sous-traitants,

- prend un certain nombre de mesures, dès lors qu'il transfère les archives dont il assure la conservation à une autre Autorité d'archivage (par exemple d'un service d'archives intermédiaires à un service d'archives définitif).

Complément : Des exemples de politiques d'archivage

Pour avoir des exemples de politiques d'archivage, se reporter sur le site de la direction des Archives de France, au chapitre 0 des Exigences-types pour la maîtrise de l'archivage électronique. Mise à jour et extension - 2008. Spécifications MoReq2 : *http://www.archivesdefrance.culture.gouv.fr/static/2085*

Les chapitres 0 prévus par les spécifications MoReq2 explicitent ce que recouvre la notion de « Records management » dans les contextes nationaux. Ce chapitre a été écrit pour le contexte français avec notamment des exemples de politiques d'archivage tant dans le secteur public (direction générale de l'Aviation civile) que dans le secteur privé (secteurs énergie, assurances, chimie).

5. Chapitre 5. Politique de sécurité à mettre en œuvre

Celle-ci doit se conformer à une politique pour la sécurité des systèmes d'information (PSSI). Cette politique de sécurité sera celle qui a été définie spécifiquement au sein de l'entreprise ou de l'établissement.

L'Autorité d'archivage doit rester responsable de la sécurité de l'ensemble des processus d'archivage, même si certaines fonctions de ces processus sont confiées à des composantes externes à l'Autorité d'archivage. Il est du ressort de l'Autorité d'archivage de faire appliquer par ces composantes, pour celles qui les concernent, les exigences de sécurité de sa politique d'archivage et d'en contrôler l'application.

Par ailleurs, la direction de l'Autorité d'archivage doit être impliquée et fournir les orientations en matière de sécurité des systèmes d'information au travers, notamment, de la validation et la diffusion d'une politique de sécurité de l'information.

L'Autorité d'archivage doit mener une analyse de risque permettant de déterminer les objectifs de sécurité propres à couvrir les risques métier de l'ensemble des processus d'archivage et les mesures de sécurité techniques et non techniques correspondantes à mettre en œuvre. Les exigences doivent être complétées et déclinées en mesures de sécurité en fonction de l'environnement réel du système d'archivage et des résultats de l'analyse de risque. En particulier la politique d'archivage doit être déclinée dans une déclaration des pratiques d'archivage

Un niveau minimal d’assurance de la sécurité offerte sur les systèmes informatiques de l'Autorité d’archivage doit être défini. Il doit au moins répondre aux objectifs de sécurité suivants :

- identification et authentification forte des utilisateurs pour l'accès au système (authentification à deux facteurs, de nature physique et/ou logique),
- gestion de sessions d’utilisation (déconnexion après un temps d’inactivité, accès aux fichiers contrôlé par rôle et nom d’utilisateur),
- protection contre les virus informatiques et toutes formes de logiciels compromettants ou non-autorisés et mises à jour des logiciels,
- gestion des comptes des utilisateurs, notamment la modification et la suppression rapide des droits d'accès,
- protection du réseau contre toute intrusion d'une personne non autorisée,
- protection du réseau afin d'assurer la confidentialité et l'intégrité des données qui y transitent,
- fonctions de traçabilité (non-répudiation et nature des actions effectuées),
- éventuellement, gestion des reprises sur erreur.

Exemple

On pourra consulter avec profit le site de l'agence nationale de la sécurité des systèmes d'information en France : *http://www.ssi.gouv.fr/*

5.1 – Les rôles de confiance

Concernant les responsabilités du niveau opérationnel, on doit distinguer au moins les rôles fonctionnels de confiance suivants :

Responsable de sécurité : le responsable de sécurité est chargé de la mise en œuvre de la politique de sécurité. Il s'assure de l'application de cette politique. Il gère les contrôles d'accès physiques aux équipements. Il est habilité à prendre connaissance des archives liées à l'activité de l'Autorité d'archivage et est chargé de l'analyse des journaux d'événements afin de détecter tout incident, anomalie, tentative de compromission, etc.

Responsable d'application : le responsable d'application est chargé de la mise en œuvre de la politique d'archivage au niveau de l'application dont il est responsable. Sa responsabilité couvre l'ensemble des fonctions rendues par cette application et des performances correspondantes.

Ingénieur système : il est chargé de la mise en route, de la configuration et de la maintenance technique des équipements informatiques. Il assure l'administration technique des systèmes et des réseaux.

Administrateur fonctionnel : il est chargé de la gestion des profils pour une application donnée.

Administrateur de base de données : il est chargé de l'administration, notamment au niveau des droits, des bases de données sur lesquelles les applications s'appuient.

Agent du service d'archivage électronique : il réalise, dans le cadre de ses attributions, l'exploitation des applications.

Attention

Concernant les rôles de confiance, les cumuls suivants sont interdits : responsable de sécurité et ingénieur système / opérateur évaluateur et tout autre rôle / ingénieur système et opérateur.

5.2 – Identification, authentification, intégrité

À partir du moment où un utilisateur aura été identifié et authentifié, il faut disposer d'un système contrôlant et limitant ses accès par rapport à ses droits. L'ensemble des utilisateurs et des droits correspondants devra être contenu dans un annuaire régulièrement mis à jour en fonction des évolutions.

Pour l'ensemble des intervenants, il est fondamental de disposer d'un dispositif permettant la parfaite identification des personnes ainsi que leur authentification. Même si un simple login mot de passe peut s'avérer suffisant pour certaines utilisations essentiellement en matière d'interrogation du SAE, pour les autres accès comme ceux directement liés à l'administration, il faudra avoir recours à des systèmes d'authentification forte lesquels garantissent que la personne identifiée est bien celle qu'elle prétend être. Le certificat électronique devra par exemple être utilisé.

Dès lors que les archives sont signalées comme confidentielles, les limites d'accès aux seules personnes habilitées doivent être gérées par le service d'archivage électronique.

Concernant les archives conservées, le contrôle d'intégrité doit avoir lieu à plusieurs niveaux du processus d' archivage :

- au niveau du transfert par le service d'archives.
- Ensuite, le contrôle d'intégrité doit pouvoir être opéré tout au long du cycle de vie des archives sans attendre des migrations éventuelles ou l'interrogation par des utilisateurs ou usagers. Des dispositifs de vérification d' intégrité basés sur l'empreinte des documents doivent ainsi être régulièrement réalisés par sondage.

5.3 –Traçabilité

Afin de constituer un ensemble de données à la fois suffisantes et cohérentes en matière de traçabilité, les opérations suivantes doivent être effectuées et validées avant la mise en place de tout service d'archivage électronique :

- Identifier les différents types d'événements à enregistrer,
- Définir les informations enregistrées pour chaque type et événement,
- Décider d'une notification ou non de l'enregistrement d'un événement au responsable de l'événement, par type d'événement et selon quelles modalités,
- Définir une fréquence a priori minimum de traitements des journaux d'événements même si toute latitude doit être laissée à ce niveau. Comme traitements, devront être entre autres analysées la recherche d'anomalies ou encore la migration des supports et des formats,
- Par rapport à ce dernier point, définir une période de conservation des journaux d'événements,
- Vérifier les dispositifs de sécurité mis en place et destinés à assurer la protection des journaux d'événements,
- Vérifier en particulier, en complément au point précédent, la procédure de sauvegarde des journaux d' événements.

Afin d'éviter toute remise en cause de l'horodatage réalisé par le service, ce dernier reposera sur un procédé conforme à l'état de l'***art*** (soit actuellement la RFC 3161 mettre un lien, tout le monde ne sait pas ce qu'est une RFC) et à tout le moins aura recours à deux sources de temps distinctes afin de dater les différentes opérations réalisées.

5.4 –Plan de continuité d'activité

Chaque composante de l'Autorité d'archivage doit disposer d'un plan de continuité d'activité permettant de répondre aux exigences de disponibilité des différentes fonctions des processus d'archivage.

Ce plan doit être testé au minimum tous les ans.

C'est ainsi que chaque entité opérant une composante de l'Autorité d'archivage doit mettre en œuvre des procédures et des moyens de remontée et de traitement des incidents, notamment au travers de la sensibilisation et de la formation de ses personnels et au travers de l'analyse des différents journaux d'événements.

Chapitre 6. Conclusion

La prise en compte de la problématique de pérennisation a des répercussions sur l'ensemble de l'organisation. Il ne s'agit donc pas de la simple mise en œuvre d'un système de plus mais bien de repenser, de faire évoluer l'organisation sous l'angle de cette approche.

Les impacts sont multiples : budgétaires, organisationnels, métiers, techniques, etc. Sans être une révolution, il s'agit d'évolutions qui selon le degré de maturité de l'organisation peuvent être longues et difficiles.

Néanmoins, nous constatons également que nous pouvons nous appuyer sur des pratiques et des usages reconnus qu'il convient d'adapter : les fondamentaux archivistiques ne sont pas remis en cause, la gestion des risques est parfaitement applicable, les bonnes pratiques informatiques sont adaptables, etc. Avec l'émergence de normes sur le sujet ce sont autant de facteurs rassurants pour une organisation.

Pour autant, les métiers d'archivistes* et de l'informatique doivent évoluer et si nous nous contentons des plans de formations nécessaires, nous risquons tout de même d'être confrontés à un sérieux déficit de compétence. La question se pose également pour les programmes d'enseignement.

Glossaire

Archivage numérique

Ensemble des actions nécessaires à la constitution et au fonctionnement d'une Archive en charge d'informations sous forme numérique.

Archives

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. Le mot archives est couramment employé dans le sens restrictif de documents ayant fait l'objet d'un archivage, par opposition aux archives courantes.

Archiviste

Professionnel chargé de la gestion d'archives.

Producteur d'archives

Personne physique ou morale, publique ou privée, qui a produit, reçu et conservé des archives dans l'exercice de son activité.

Bibliographie

• BANAT-BERGER F., HUC C., DUPLOUY L., L'Archivage numérique à long terme, les débuts de la maturité? Paris, La Documentation française, 2009.

Webographie

http://public.ccsds.org/publications/archive/650x0b1(F).pdf