



Étude et réalisation d'un prototype d'archivage électronique appliqué aux collectivités territoriales

Julien Montel

► To cite this version:

Julien Montel. Étude et réalisation d'un prototype d'archivage électronique appliqué aux collectivités territoriales. Recherche d'information [cs.IR]. 2016. dumas-01798631

HAL Id: dumas-01798631

<https://dumas.ccsd.cnrs.fr/dumas-01798631>

Submitted on 23 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONSERVATOIRE NATIONAL DES ARTS ET MÉTIERS

CENTRE RÉGIONAL ASSOCIÉ DE RENNES

MÉMOIRE

présenté en vue d'obtenir

le DIPLÔME D'INGÉNIEUR CNAM

SPÉCIALITÉ : Informatique

OPTION : Architecture et ingénierie des systèmes et des logiciels (AISL)

par

Julien MONTEL

**Étude et réalisation d'un prototype d'archivage électronique appliquée aux
collectivités territoriales**

Soutenu le 21 janvier 2016

JURY

PRÉSIDENT : M. Yann POLLET (CNAM Paris)

MEMBRES : M. Charles PRÉAUX (CNAM Bretagne)
M. Thierry FORVEILLE (Rennes Métropole)
M. David CHARIL (Rennes Métropole)

Remerciements

Tout d'abord, je remercie mon tuteur CNAM, Charles Préaux pour son soutien et ses précieux conseils.

Je remercie également mes deux tuteurs et collègues, Thierry Forveille et David Charil, autant pour leurs conseils que leur disponibilité.

Un grand merci aux collègues de la Direction des systèmes d'Information, des Archives Municipales ou encore de la Direction de la Commande Publique qui m'ont apporté aussi bien des conseils, du soutien moral ou encore des précisions sur le fonctionnement des deux institutions pour lesquelles je travaille.

Je remercie aussi Patrick Percot du Centre de Gestion du Morbihan pour les réponses qu'il m'a apportées et pour le code source qu'il m'a fourni.

Merci également à Nathalie Marin qui m'aura soutenu dans ce travail.

Enfin, je tiens à remercier ma femme Aurélie ainsi que mes enfants Olivia et Samuel pour leur soutien, leur compréhension et leur patience.

Liste des abréviations

- **AAF** : Association des Archivistes Français
- **ADAE** : Agence pour le Développement de l'Administration Électronique
- **ADELE** : ADministration ÉLectronique
- **ADULLACT** : Association des Développeurs et Utilisateurs de Logiciels Libres pour les Administrations et les Collectivités Territoriales
- **AFNOR** : Association Française de Normalisation
- **AFS** : Archives Fédérales Suisses
- **AIP** : Archival Information Package (Paquet d'Information Archivé)
- **ANSSI** : Agence Nationale pour la Sécurité des Systèmes d'Information
- **API** : Application Protocol Interface
- **art.** : Article (*législatif*)
- **ASCII** : American Standard Code for Information Interchange
- **BnF** : Bibliothèque Nationale de France
- **Bull.** : Bulletin (*législatif*)
- **CAP** : Création, Architecture et Patrimoine
- **Cass.** : Chambre de Cassation (*législatif*)
- **C. civ.** : Code Civil (*législatif*)
- **C. patr.** : Code du Patrimoine (*législatif*)
- **CCAS** : Centre Communal d'Action Social
- **CCSDS** : Consultative Committee for Space Data Systems
- **CEN** : Comité Européen de Normalisation
- **cf.** : Confer
- **CFN** : Coffre-Fort Numérique
- **CGCT** : Code Général des Collectivités Territoriales
- **CIA** : Conseil International des Archives
- **Civ. 2** : 2^{ème} Chambre Civile (*législatif*)
- **CL** : Contrôle de Légalité
- **CMIS** : Content Management Interoperability Services
- **CNAM** : Conservatoire National des Arts et Métiers
- **CNES** : Centre National d'Études Spatiales
- **CSV** : Comma-Separated Values
- **D.** : Décret (*législatif*)
- **DAF** : Direction des Archives de France
- **DGME** : Direction Générale de la Modernisation de l'Etat (*ex ADAE*)
- **DG** : Direction Générale
- **DGCult** : Direction Générale de la Culture
- **DGFIP** : Direction Générale des Finances Publiques
- **DGS** : Directeur Générale des Services
- **DISIC** : Direction Interministérielle des Systèmes d'Information et de Communication

- **DIP** : Dissemination Information Package (Paquet d'Information à Diffuser)
- **DLM Forum Foundation** : Document Lifecycle Management Forum Foundation
- **DSI** : Direction des Systèmes d'Information
- **DTD** : Document Type Definition (Définition de Type de Document)
- **DUA** : Durée d'Utilité Administrative
- **DUC** : Durée d'Utilité Courante
- **EAC-CPF** : Encoded Archival Context – Corporate Bodies, Persons and Families (Contexte Archivistique Encodé – Collectivités, Personnes et Familles)
- **EAD** : Encoded Archival Description (Description Archivistique Encodée)
- **ERMS** : Electronic Records Management System (*voir SAE en français*)
- **GIS-SPADON** : Groupement d'Intérêt Scientifique sur les Supports Pérennes d'Archivage des DONnées Numériques
- **IDE** : Integrated Development Environment
- **IHM** : Interface Homme Machine
- **ISAAAR (CPF)** : International Standard Archival Authority Record for Corporate Bodies, Persons and Families (Norme Internationale sur les Notices d'Autorité utilisées pour les Archives relatives aux Collectivités, aux Personnes et aux Familles)
- **ISAD (G)** : International Standard Archival Description-General (Norme Générale et Internationale de Description des Archives)
- **ISO** : International Organization for Standardization (Organisation Internationale de Normalisation)
- **LNE** : Laboratoire National de métrologie et d'Essais
- **Maptam** : Modernisation de l'Action Publique Territoriale et d'Affirmation des Métropoles
- **MDS** : Message Digest
- **MEDONA** : Modélisation des Échanges de DONnées pour l'Archivage
- **METS** : Metadata Encoding and Transmission Standard
- **Moreq** : Model Requirements for the Management of Electronic Records (Exigences Types pour la Maîtrise de l'Archivage Électronique)
- **NAS** : Network Attached Storage (Serveur de Stockage en Réseau)
- **OAIS** : Open Archival Information System (Système Ouvert d'Archivage d'Information)
- **OASIS** : Organization for the Advancement of Structured Information Standards
- **OMG** : Object Management Group
- **ord.** : Ordinance (*législatif*)
- **PAIS** : Producer-Archive Interface Specification
- **PAIMAS** : Producer Archive Interface Methodology Abstract Standard
- **PDF** : Portable Document Format
- **PES v2** : Protocole d'Échange Standard version 2
- **PIAF** : Portail International Archivistique Francophone
- **PIN** : Pérennisation de l'Information Numérique

- **PREMIS** : PREservation Metadata: Implementation Strategies
- **RAID** : Redundant Arrays of Inexpensive Disks (Regroupement Redondant de Disques Indépendants)
- **RelaxNG** : Regular Language for XML Next Generation
- **REST** : Representational State Transfer
- **RGAA** : Référentiel Général d'Accessibilité pour les Administrations
- **RGI** : Référentiel Général d'Interopérabilité
- **RGS** : Référentiel Général de Sécurité
- **RM** : Record Management
- **SAA** : Society of American Archivists (Société des Archivistes Américains)
- **SaaS** : Software As A Service
- **SAE** : Système d'Archivage Électronique (*voir ERMS en anglais*)
- **SAN** : Storage Area Network (Réseau de Stockage)
- **SEDA** : Standard d'Échange de Données pour l'Archivage
- **SGBDR** : Système de Gestion de Bases de Données Relationnelles
- **SGDA** : Système de Gestion des Documents d'Activité
- **SGMAP** : Secrétariat Général pour la Modernisation de l'Action Publique
- **SHA** : Secure Hash Algorithm
- **SIAF** : Service Interministériel des Archives de France
- **SIARD** : Software Independent Archiving of Relational Databases
- **SIP** : Submission Information Package (Paquet d'Information à Verser)
- **SMTP** : Simple Mail Transfer Protocol
- **SOA** : Service Oriented Architecture
- **SOAP** : Simple Object Access Protocol
- **SQL** : Structured Query Language
- **UML** : Unified Modeling Language
- **UTF-8** : Universal character set Transformation Format
- **VITAM** : Valeurs Immatérielles Transférées aux Archives pour Mémoire
- **W3C** : World Wide Web Consortium
- **WORM** : Write Once Read Many
- **XML** : eXtensible Markup Language
- **XSD** : Schéma XML

Glossaire

- **Collecte** : Une des principales missions d'un service d'archive, qui consiste à rechercher et recueillir auprès des producteurs les différents documents à archiver.
- **Communicabilité** : Possibilité de communiquer au public un document archivé en fonction de critères juridiques et réglementaires.
- **Communication** : Mission d'un service d'archive consistant en la mise à disposition de documents archivés communicables au public et aux administrations.
- **Conservation définitive** : Opération qui consiste à conserver les documents qui ne sont pas destinés à être éliminés.
- **Date de communicabilité** : Date à partir de laquelle un document devient communicable, c'est-à-dire accessible au public.
- **Document** : Tout écrit ou enregistrement considéré comme une unité (*extrait norme ISO 15489*).
- **Dossier** : Ensemble de documents constitués de manière organique ou logique.
- **Élimination** : Procédure réglementée qui vise à supprimer ou retirer du système d'archivage un ou plusieurs documents archivés.
- **Évaluation** : Fonction de l'archiviste qui consiste à déterminer l'utilité administrative, l'intérêt historique et le sort final des documents.
- **Instrument de recherche** : Outil papier ou informatique permettant la recherche d'archives.
- **Métadonnées** : Données décrivant le contexte, le contenu, la structure des documents ainsi que leur gestion dans le temps.
- **Pièce** : Plus petite unité intellectuelle d'archives, par exemple une lettre, un mémoire, une photo...
- **Plan de classement** : Système fixant l'organisation des archives courantes, intermédiaires et définitives.
- **Record Management** : Gestion des documents d'archives afin d'organiser efficacement et systématiquement la création, la réception, l'utilisation et le sort des documents d'archives. On dit également archivage managérial.
- **Sort Final** : Désigne le traitement final des documents.
- **Tri** : Opération qui consiste à séparer les documents dont le sort final sera la conservation de ceux qui devront être éliminés.
- **Versement** : Opération qui consiste à transmettre des documents à archiver à un service (ou système) d'archivage.
- **Valeur probante** : Qualité d'un document lui permettant de faire valeur de preuve.

Table des matières

Introduction.....	10
1. Contexte	11
2. Présentation du sujet	12
a. Gestion Electronique Documentaire	13
b. Archivage légal	13
c. Archivage patrimonial	13
3. Objectifs	14
4. Réduction du périmètre	15
5. Gestion de projet.....	15
a. Mon rôle dans le projet	15
b. Planning prévisionnel	16
6. Analyse des risques du projet	17
a. Liste des risques.....	17
b. Cotation des risques	17
c. Maîtrise des risques ou recommandations	18
7. Annonce du plan	20
I Présentation de Rennes Métropole / Ville de Rennes et leurs environnements dans le projet.....	21
I.1 Rennes Métropole	21
I.1.1 Domaines d'intervention	22
I.1.2 Instances et rôles	24
I.2 La Ville de Rennes.....	25
I.2.1 Quelques chiffres	25
I.2.2 Instances et rôles	25
I.3 Direction des Systèmes d'Information	26
I.3.1 Le Service Etudes et Projets	28
I.3.2 Le Service Infrastructure.....	29
I.3.3 Le Service Poste de Travail.....	30
I.3.4 Le Service Assistance Administrative	31
I.4 Service des Archives Municipales	31

I.5	Mégalis Bretagne	32
II	État de l'art	33
II.1	Présentation des archives	33
II.1.1	Finalités des archives	34
II.1.2	Actions sur une archive	35
II.1.3	Théorie des 3 âges	35
II.1.4	Record Management	37
II.2	Textes de référence	38
II.2.1	Cadre juridique et réglementaire	38
II.2.2	Cadre normatif et standards	41
II.3	Formats et supports de stockage préconisés pour la pérennisation	55
II.3.1	Encodage	55
II.3.2	Préserver les documents en fonction des formats utilisés	56
II.3.3	Supports de stockage	59
II.3.4	Migrations	63
III	Spécifications du démonstrateur	64
III.1	SAE	64
III.1.1	Distinctions SAE, GED et CFN	64
III.2	Spécifications fonctionnelles et techniques	67
III.2.1	Spécifications fonctionnelles	70
III.2.2	Synthèse des exigences	73
III.3	Architecture type de la solution	75
III.4	Étude de marché	77
III.4.1	Solutions étudiées	77
III.4.2	Comparatif des différentes solutions étudiées	78
III.4.3	Choix de la solution retenue pour la maquette	80
III.4.4	Autres solutions	80
IV	Réalisation et mise en œuvre du prototype	82
IV.1	Installation de la solution	82
IV.1.1	Installation des composants	82
IV.1.2	Architecture d'as@lae	87

IV.2 Paramétrages as@lae.....	88
IV.2.1 Accès à l'application	88
IV.2.2 Paramétrages fonctionnels	89
IV.3 Tests depuis interfaces as@lae	94
IV.3.1 Versement fichier PDF version 1.4 et XML.....	95
IV.3.2 Versement fichier PDF version 1.5 et docx	99
IV.3.3 Versement fichier PDF version 1.2	99
IV.3.4 Autres tests effectués	100
IV.4 Test versement automatique depuis la GED Alfresco.....	100
IV.4.1 Ajouts de métadonnées sur la GED Alfresco.....	101
IV.4.2 Connecteur pour transférer le document de la GED au SAE	103
IV.5 Évaluation du prototype.....	105
Conclusion	108
Bibliographie	109
1. Sources et citations	109
2. Lectures pour aller plus loin	113
Liste des figures.....	114
Liste des tableaux.....	116
Liste des annexes.....	117

Introduction

Le numérique ne cesse de prendre de l'ampleur dans nos sociétés et les activités professionnelles n'échappent pas à cette règle. Des PME artisanales aux grandes multinationales, en passant par les collectivités territoriales et les autres institutions, la plupart des échanges et des décisions sont progressivement dématérialisés.

En effet les devis, factures, bons de commande, bulletins de salaire, contrats et autres arrêtés sont générés par des logiciels de bureautique ou logiciels métiers. Même dans le domaine culturel, les fichiers multimédias tels que les photos, les vidéos ou encore les enregistrements audio sont également dématérialisés.

Cette tendance forte implique un changement dans les manières de penser la conservation à plus ou moins long terme des documents, à des fins de sécurisation réglementaire. En effet, certains documents ont une valeur juridique ou contractuelle et sont soumis à un délai de prescription. En fonction du type de document généré, cette durée peut varier. Pour une entreprise par exemple, les pièces comptables doivent être conservées pendant 10 ans, un contrat d'acquisition de biens immobiliers sera lui conservé 30 ans... [1]¹.

De plus, certains documents présentent un intérêt historique et par conséquent leur conservation permet de préserver la mémoire de l'humanité.

Cette conservation relève des compétences des services d'archives de l'entreprise ou de la collectivité. Ces services sont en pleine mutation du fait de ce bouleversement numérique. Ils doivent en effet continuer à gérer des archives papier comme ils le faisaient auparavant mais également intégrer la gestion des documents électroniques. Ce secteur encore jeune est l'objet de beaucoup de débats afin de pérenniser les données générées par les entreprises privées et publiques et voit sa réglementation évoluer régulièrement.

Avant d'aller plus avant, il est nécessaire de définir ce que nous entendons par « archives » et « archivage ».

Les archives ne sont pas des sauvegardes ou des moyens de stockage, ces deux termes recouvrant une notion de copie uniquement. Selon la définition du Code du Patrimoine (**c. patr., art. L211-1**), il s'agit « [...] des documents, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité ». « L'archivage est le transfert de documents qui ont cessé d'être d'utilité courante dans un lieu de stockage ou un service d'archives compétent pour les recevoir » [2]². L'Association Française de Normalisation (AFNOR) propose la définition de l'archivage comme étant « l'ensemble des actions, des outils et des méthodes mises en œuvre pour conserver à moyen

¹ Premier Ministre : Quels sont les délais de conservation des documents pour les entreprises ?

² Direction des Archives de France : Dictionnaire de terminologie archivistique

ou long terme des informations dans le but de les exploiter ».

Il est utile de préciser également que le cycle de vie des archives est constitué de trois étapes, **les archives courantes, les archives intermédiaires et les archives définitives**. Les archives courantes concernent les documents fréquemment consultés ou modifiés. Les archives intermédiaires ne subissent plus de modification mais sont encore consultées ponctuellement. Elles sont donc conservées pendant toute leur durée d'utilité administrative (*DUA*). Enfin, les archives définitives sont conservées sans limitation de durée, généralement pour des raisons historiques ou patrimoniales [3]¹.

Un service d'archives publiques a pour mission de collecter, classer, conserver et communiquer les différents documents créés par les services publics.

1. Contexte

Je travaille depuis 2005 à la Direction des Systèmes d'Information (*DSI*) de Rennes Métropole / Ville de Rennes. La collectivité génère de plus en plus de documents dématérialisés. Depuis quelques années et entre autre depuis l'ancien plan ADELE (*ADministration ÉLectronique, 2004-2007*), maintenant terminé et repris par d'autres services, l'État incite les collectivités locales à dématérialiser leurs différentes procédures. Cette dématérialisation est d'une complexité qui n'avait pas été bien mesurée et entraîne des impacts périphériques.

Les différents documents générés, émis, reçus doivent pouvoir être accédés, conservés, archivés à différentes fins et surtout ne doivent pas perdre de leur valeur administrative ou juridique. Cette dématérialisation entraîne une modification dans la conservation de ces documents tout au long de leur durée de vie.

En prenant l'exemple d'une passation d'un marché public, la collectivité élabore tout d'abord sous format bureautique des documents précisant la nature, l'objectif dudit marché. Ces documents sont actuellement disponibles sur les espaces partagés de stockage fournis par la DSI. Ces documents constituent un cadre de réponse pour les entreprises qui souhaitent postuler au marché. De plus en plus de réponses se font de manière électronique.

La réglementation impose d'ailleurs aux entreprises de répondre sous format électronique pour les marchés publics informatiques d'un montant supérieur à 90 000 € HT. Rennes Métropole ne dispose pas actuellement de certificat électronique permettant de signer les documents dématérialisés. Il est par conséquent demandé aux entreprises, et

¹ Locarchives : Qu'est-ce qu'une archive ?

surtout à celle qui sera retenue, d'une part de répondre électroniquement au marché, la réponse contenant donc un certificat électronique, mais également de matérialiser leur réponse et d'y apposer une signature manuscrite.

Une autre raison de la matérialisation est que le transfert des marchés sous forme dématérialisée n'est pas actuellement mis en place entre la collectivité territoriale et le contrôle de légalité (*CL*), organisme d'État en charge de la vérification de conformité du marché.

Toutefois, une fois le marché devenu exécutoire, les flux comptables sont et devront obligatoirement être dématérialisés. En effet, en application du Code Général des Collectivités Territoriales (*CGCT*) (**CGCT, art.1617-23**), les collectivités territoriales doivent depuis le 1^{er} janvier 2015 transmettre de manière dématérialisée les flux comptable aux trésoreries (bordereaux, mandats, titres).

Depuis cette échéance, elles doivent avoir adopté le nouveau format d'échange, le « Protocole d'Échange Standard version 2 » (*PES v2*) en abandonnant l'ancien protocole Indigo. Ce changement s'est effectué pour les collectivités avec l'aide de la Direction générale des Finances publiques (*DGFiP*), ainsi que de l'éditeur logiciel du logiciel financier (Coriolis de Bull pour la Ville de Rennes et Civil Net Finance de Ciril pour Rennes Métropole). À terme, le *PES v2* prévoit que soient dématérialisés les flux comptables et également les pièces justificatives (factures, fiches de paie, délibérations, arrêtés et décisions, pièces de passation de marchés...).

Outre les passations de marchés publics et les flux comptables, les délibérations et autres actes sont également en cours de dématérialisation.

L'exemple décrit ci-dessus montre que les échanges et par conséquent les documents générés ou reçus par la collectivité seront de plus en plus dématérialisés. Il convient donc de penser la manière de stocker, de conserver, de préserver et d'accéder à ces documents durant leur durée de vie.

Entre leur création, leurs éventuelles modifications, la durée pendant laquelle ils doivent être conservés à des fins légales, et enfin leur conservation à plus long terme dans un but historique, le simple stockage de fichiers n'est pas une solution satisfaisante.

2. Présentation du sujet

Rennes Métropole se doit donc de mettre en œuvre une solution d'archivage électronique. Ce vaste sujet qui sera en partie traité dans ce document doit être repositionné dans sa globalité. Il est en effet articulé autour de trois grandes composantes liées entre elles. Il s'agit de la mise en œuvre d'un référentiel documentaire, de l'archivage légal (à valeur probante) et enfin de l'archivage patrimonial.

a. Gestion Électronique Documentaire

La mise en place d'un espace documentaire est un projet en cours, basé sur une Gestion Électronique Documentaire (*GED*) construite avec l'outil Alfresco Enterprise. Cette GED permettra entre autre de gérer les versions de documents créés ou reçus par la collectivité.

Ce projet de GED contient également un parapheur électronique (i-parapheur de l'Adullact). Ce parapheur permettra à terme les signatures électroniques. Toutefois, il n'est pas prévu que la collectivité dispose de certificat électronique dans un premier temps. Par conséquent, il sera tout d'abord utilisé à la manière d'un workflow, permettant d'apposer uniquement des visas, la valeur contractuelle ou juridique restant la signature manuscrite de l'élu compétent.

Enfin, ce volet de GED est étroitement lié avec un autre projet. Il s'agit de la mise en place d'espaces collaboratifs, afin de faciliter l'échange de documents, le travail inter-directions lors de la mise en place de futurs projets ou encore la constitution de référentiels documentaires fiables et uniques.

b. Archivage légal

Le deuxième point traite d'un système d'archivage légal (ou intermédiaire). Ce projet, également en cours, est piloté par David Charil, chef de projet au sein du service Études et Projets de la Direction des Systèmes d'Information.

La solution est basée sur un Système d'Archivage Électronique (*SAE*) intermédiaire, fourni par un prestataire externe, Mégalis Bretagne. Il s'agit d'*as@lae, SAE Open Source développé par l'Adullact*¹, hébergé chez Navaho, tiers archiveur agréé par le Service Interministériel des Archives de France (*SIAF*).

Mégalis est un syndicat mixte de coopération territoriale ayant entre autres pour compétence le développement de l'administration électronique. Mégalis doit fournir un *SAE* permettant de stocker via des connecteurs les documents qui ont une DUA. C'est-à-dire les documents dont l'accès est encore nécessaire par la collectivité, et dont la valeur juridique est encore valide.

c. Archivage patrimonial

Le troisième et dernier volet de ce vaste projet concerne l'archivage patrimonial. Cette mise

¹ Association des Développeurs et Utilisateurs de Logiciels Libres pour les Administrations et les Collectivités Territoriales

en œuvre interviendra ultérieurement, mais Rennes Métropole doit dès à présent en aborder la réflexion. À cet effet, il apparaît nécessaire de réaliser dès aujourd’hui un état de l’art ainsi que la qualification d'une solution par la réalisation d'un prototype.

Il s’agira à terme de mettre en place un SAE pour les archives patrimoniales, archives destinées à être conservées indéfiniment. C'est sur ce troisième projet que seront axées la réflexion et l'analyse de ce document.

Le schéma ci-dessous illustre l'ensemble des flux du projet :

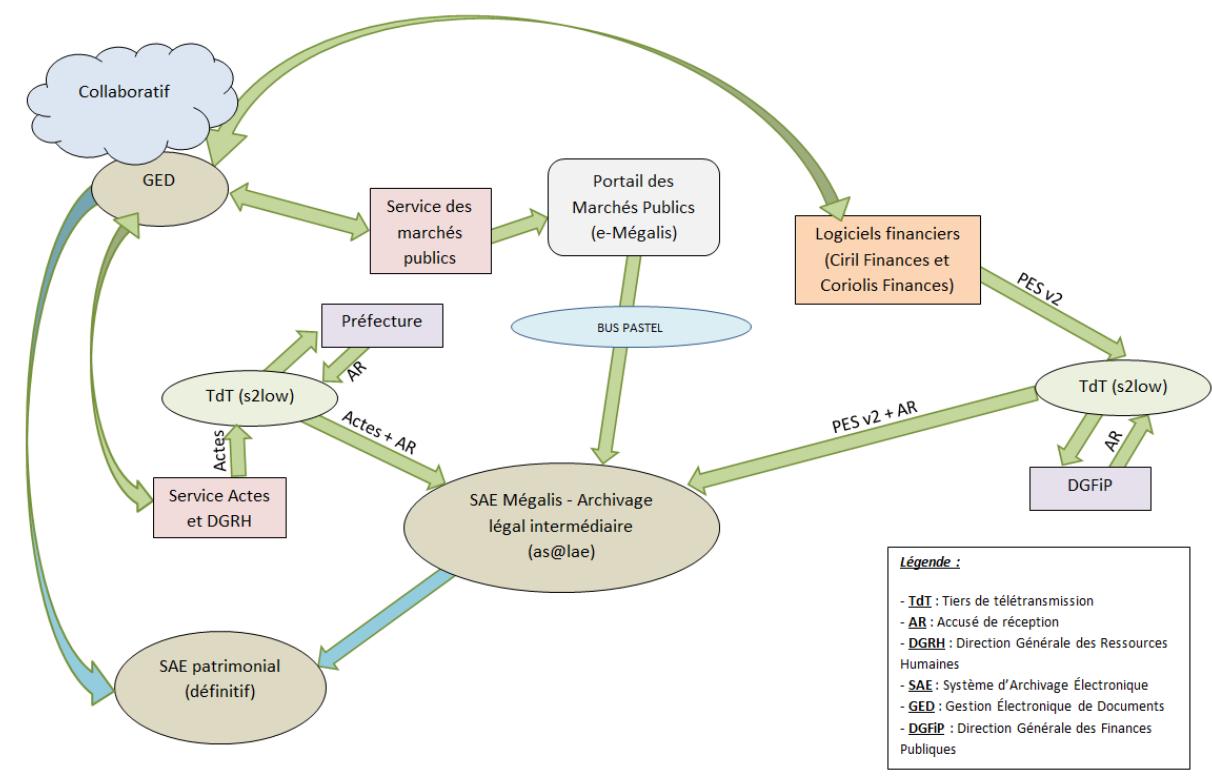


Figure 1 : Schéma illustrant les différentes briques du projet dans sa globalité

3. Objectifs

Le but de la présente analyse est d'effectuer un état de l'art des normes et solutions techniques existantes dans le domaine de l'archivage électronique, de présenter un panorama de ces solutions et de définir les bons critères de choix pour en retenir une. Enfin, il est également prévu de spécifier, concevoir, réaliser et mettre en œuvre un prototype d'une telle solution, en créant également des connecteurs entre le système de **GED Alfresco** et celui du **SAE d'archivage intermédiaire**.

À travers ces différents objectifs, la finalité de ce document est d'être une aide au choix d'un futur Système d'Archivage Électronique patrimonial pour les collectivités mutualisées

Rennes Métropole / Ville de Rennes.

4. Réduction du périmètre

Étant donné l'ampleur du domaine des SAE et de l'archivage électronique de manière générale, le présent document se limitera à l'étude, la spécification et la mise en œuvre d'un prototype de SAE dans le cadre des collectivités territoriales et plus particulièrement pour les collectivités Ville de Rennes / Rennes Métropole dans un contexte de mutualisation. Cette maquette ne couvrira pas toutes les fonctionnalités que l'on peut attendre d'un tel système, mais en évoquera au moins l'existence et la nécessité. D'un point de vue technique, elle devra permettre de transmettre depuis l'espace de la GED ou depuis le SAE intermédiaire fourni par Mégalis, les documents dans des formats définis et selon les dernières normes en matière d'archivage numérique.

La partie traitant de l'état de l'art de l'archivage électronique n'abordera pas tous les aspects liés à l'archivage. Ne sera notamment pas abordée dans son intégralité la législation en vigueur, qui touche plus au métier d'archiviste. Le présent document entend plutôt être une aide au choix d'une solution à la frontière entre les deux domaines que sont l'informatique et l'archivistique.

5. Gestion de projet

Ce vaste projet est dirigé globalement par Thierry Forveille, responsable du Service Études et Projets de la DSI. Les projets de GED et d'archivage légal sont du ressort de David Charil, chef de projets à la DSI.

a. Mon rôle dans le projet

Étant donné la dématérialisation des processus et l'obligation liée à l'archivage, Rennes Métropole doit dès maintenant réaliser un état de l'art et la qualification d'une solution sur l'archivage électronique définitif.

Mon rôle sera la conduite de ce projet dans l'équipe projet composée de Thierry Forveille, David Charil et Romain Joulia, Directeur des Archives de Rennes.

b. Planning prévisionnel

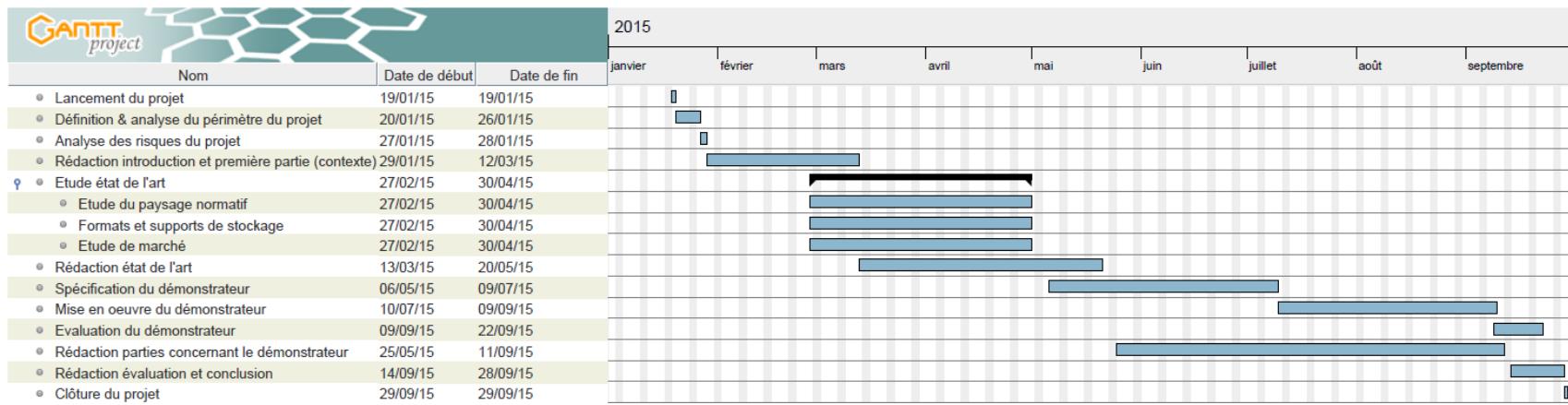


Figure 2 : Diagramme de GANTT - Planning prévisionnel du projet de mémoire

6. Analyse des risques du projet

La problématique d'archivage électronique est transverse à différents acteurs internes et externes à la collectivité, ainsi qu'à différents projets en cours ou qui seront prochainement lancés.

En effet, tous les services sont impactés par l'archivage, mais certains services seront plus impliqués dans les projets transversaux.

Au niveau de la DSI, le projet global est transverse au projet de GED, au projet d'espace collaboratif, au projet d'archivage intermédiaire qui sera fourni par le prestataire externe Mégalis Bretagne, ainsi qu'au projet de refonte du système d'information des archives municipales.

Au regard des enjeux, il nous a semblé nécessaire d'effectuer une analyse des risques, ne serait-ce que succincte, afin de réduire les possibilités que le projet n'aboutisse pas.

a. Liste des risques

Nous avons d'abord répertorié certains des risques susceptibles de survenir lors de ce projet. En voici la liste :

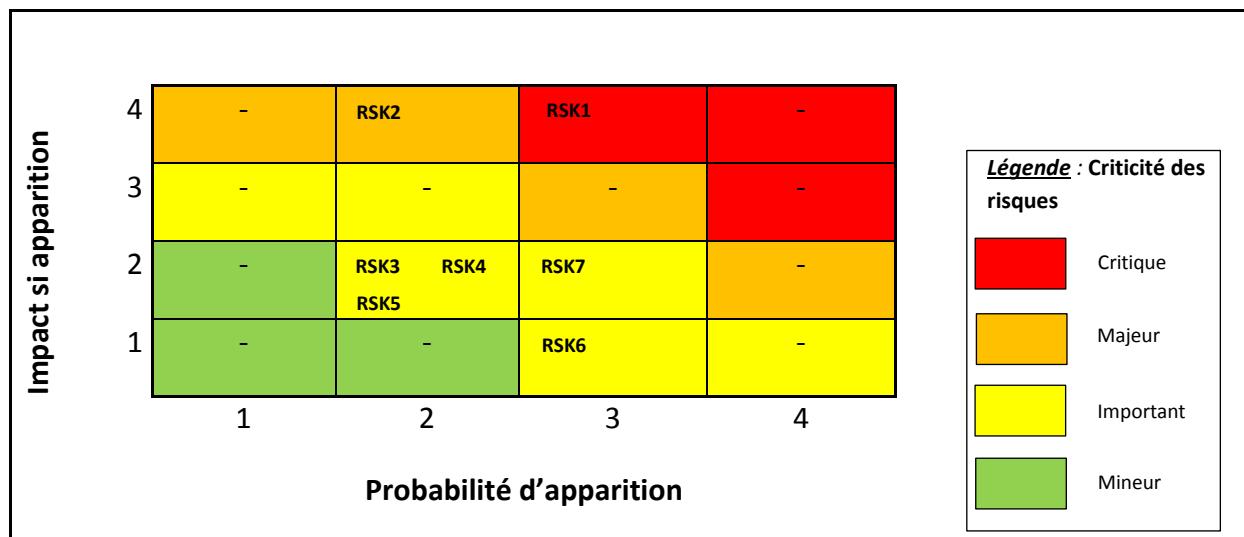
- **RSK1** : Retard de livraison du SAE intermédiaire de Mégalis.
- **RSK2** : Retard dans la mise en place du projet de GED.
- **RSK3** : Non disponibilité de l'équipe système pour fournir une machine virtuelle (pour mise en œuvre de la maquette).
- **RSK4** : Espace disque non disponible pour mise en œuvre de la maquette.
- **RSK5** : Non disponibilité des acteurs du service des archives.
- **RSK6** : Évolution des solutions de SAE.
- **RSK7** : Évolution de la législation en matière d'archivage définitif pour les collectivités.

b. Cotation des risques

Nous avons choisi de coter le risque en lui affectant une probabilité d'apparition, notée de 1 à 4, 1 correspondant à la probabilité d'apparition la plus basse, proche de 0, et 4 correspondant à une probabilité d'apparition importante.

Le risque est coté également avec une note positionnée sur l'impact qu'occasionnerait l'apparition du risque. Cet impact a également une plage de 1 à 4, de l'impact le moins grave à l'impact le plus grave.

Voici ci-dessous la matrice correspondant à la cotation des risques répertoriés ci-avant :

Tableau I : Matrice de cotation des risques

Cette matrice des risques fait ressortir les points suivants :

Un risque critique, qui est à traiter en priorité dans le plan d'action, un risque majeur qui doit être également minimisé le plus tôt possible, et cinq risques importants, qui peuvent être acceptés, mais pour lesquels il est tout de même préférable de prendre les dispositions nécessaires afin de les réduire.

c. Maîtrise des risques ou recommandations

- **RSK1** : Retard de livraison du SAE intermédiaire de Mégalis.

On constate que le risque critique est celui qui serait dû au retard de livraison du SAE intermédiaire par Mégalis.

En effet, la maquette à réaliser devra permettre de récupérer des documents depuis le SAE Mégalis, via des Web Services par exemple.

Étant donné qu'il s'agit d'un prestataire externe, nous n'avons que peu de maîtrise afin de réduire cette criticité.

Par contre, afin de réduire l'apparition de ce risque, il est possible tout d'abord que le Chef de Projet les relance fréquemment. Il est également conseillé de prendre contact avec d'autres collectivités qui ont choisi cette solution afin de se renseigner.

De plus des collectivités territoriales, telles que Brest Métropole Océane (*BMO*) et le Centre de Gestion 56 (*CDG56*), ont déjà mis en place la solution du SAE Mégalis. Il serait donc intéressant de les rencontrer pour un retour d'expérience.

- **RSK2** : Retard dans la mise en place du projet de GED.

Ce risque est coté comme étant majeur. Il sera également un axe principal du futur prototype. Les documents devront partir de cet espace vers le SAE Mégalis ou le SAE définitif, par conséquent la maquette.

Le risque de retard sur cette étape correspond surtout à un manque de disponibilité du chef de projet GED car il gère par ailleurs d'autres projets. Il peut être minimisé en débutant la maquette le plus tôt possible, laissant par conséquent une marge plus importante au niveau calendaire. Ce risque peut également être maîtrisé avec une bonne compréhension des mécanismes de la nouvelle version de la GED par l'ensemble des acteurs.

- **RSK3** : Non disponibilité de l'équipe système pour fournir une machine virtuelle (pour mise en œuvre de la maquette).

La non-disponibilité des techniciens système pourrait entraîner un retard de livraison d'une plateforme virtuelle utilisée pour la maquette. Ce risque peut être réduit en effectuant la demande de cette machine virtuelle le plus en amont possible ou bien en la réalisant sur un PC puissant plutôt que sur une ferme de serveurs virtualisés.

- **RSK4** : Espace disque non disponible pour mise en œuvre de la maquette.

Ce risque aurait également un impact au niveau du prototype. Il peut être minimisé en effectuant la demande d'espace le plus tôt possible, ou en revoyant à la baisse le besoin.

- **RSK5** : Non disponibilité des acteurs du service des archives.

Les agents travaillant au service des archives connaissent le processus métier d'archivage, les règles de durée de conservation des documents, quels documents sont à archiver ou pas. Une non-disponibilité de leur part aurait pour conséquence la mise en œuvre d'une maquette qui ne respecterait pas les fondements du métier d'archiviste. Par conséquent, ce risque peut être diminué en posant les questions le plus rapidement possible.

- **RSK6** : Évolution des solutions de SAE.

L'archivage électronique est un sujet encore très jeune. Les architectures et technologies permettant la mise en place d'un SAE évoluent donc fréquemment.

Il faudrait éviter de proposer une architecture ou de se baser sur une technologie désuète ou qui n'est plus d'actualité pour la mise en œuvre de la maquette. Ce risque sera minimisé en effectuant une veille le plus en amont possible. En matière de veille, l'Adullact effectue début 2015 un « Tour de France Archivage » dans les collectivités territoriales afin de présenter leur solution libre de SAE électronique, As@lae. Il sera donc intéressant d'y participer afin de prendre connaissance de leur solution.

- **RSK7** : Évolution de la législation en matière d'archivage définitif pour les collectivités.

Comme pour le risque précédent, le domaine de l'archivage électronique étant récent, la législation qui l'encadre peut également évoluer prochainement. Cette évolution peut toucher en particulier les obligations en matière d'archivage définitif pour les collectivités territoriales. À l'heure où sont rédigées ces lignes, les collectivités territoriales ne peuvent pas externaliser leur archivage définitif, écartant par là même des solutions basées sur du Cloud externe, du fait de l'hébergement hors du territoire de la collectivité. Une évolution de la législation sur ce domaine pourrait faire inverser le choix d'une solution. Là encore, pour réduire ce risque, une veille législative est nécessaire.

Nous avons abordés les principaux risques qu'il nous a semblé important d'identifier en amont du projet afin de les minimiser le plus possible. Nous allons maintenant présenter notre logique de progression.

7. Annonce du plan

Nous effectuerons en première partie de ce document une présentation des collectivités Rennes Métropole et Ville de Rennes, ainsi que de leurs environnements dans le cadre du projet.

Ensuite, nous dresserons un état de l'art de l'archivage numérique à l'heure actuelle, en particulier comment répondre aux problématiques touchant la conservation de l'information numérique, quelles normes, standards, supports de stockage ou formats sont préconisés dans ce cadre.

Après avoir effectué ce panorama, le démonstrateur à développer sera spécifié, et enfin sa réalisation et sa mise en œuvre seront décrites dans la quatrième partie.

I Présentation de Rennes Métropole / Ville de Rennes et leurs environnements dans le projet

L'objectif de ce premier chapitre est de présenter les deux collectivités pour lesquelles je travaille. Nous ferons également une présentation de la Direction des Systèmes d'Information au sein de laquelle j'exerce mon activité, ainsi que du service des Archives Municipales et des autres interlocuteurs qui interviennent sur ce sujet. Il convient de signaler qu'il est prévu une restructuration des organigrammes des collectivités Ville de Rennes et Rennes Métropole courant 2015. Cette présentation ne tiendra pas compte de ce futur changement car elle est sans impact sur le projet.

I.1 Rennes Métropole

Rennes Métropole est depuis le 1^{er} janvier 2015 une des 14 métropoles du territoire français. Ce statut succède à celui de communauté d'agglomération suite à la loi Maptam (**L. n°2014-58, 27 janv. 2014** ou loi de **Modernisation de l'Action Publique Territoriale et d'Affirmation des Métropoles**).

Rennes Métropole est composée actuellement de 43 communes et compte plus de 420 000 habitants (425 745 au 1^{er} janvier 2014 selon l'INSEE).

Son budget 2014 s'élevait à 642,45 M€ et 829,7 M€ en 2015 (dont 463,88 M€ d'investissement pour la construction de la deuxième ligne de métro).

Ci-dessous la carte présentant les communes membres de Rennes Métropole :

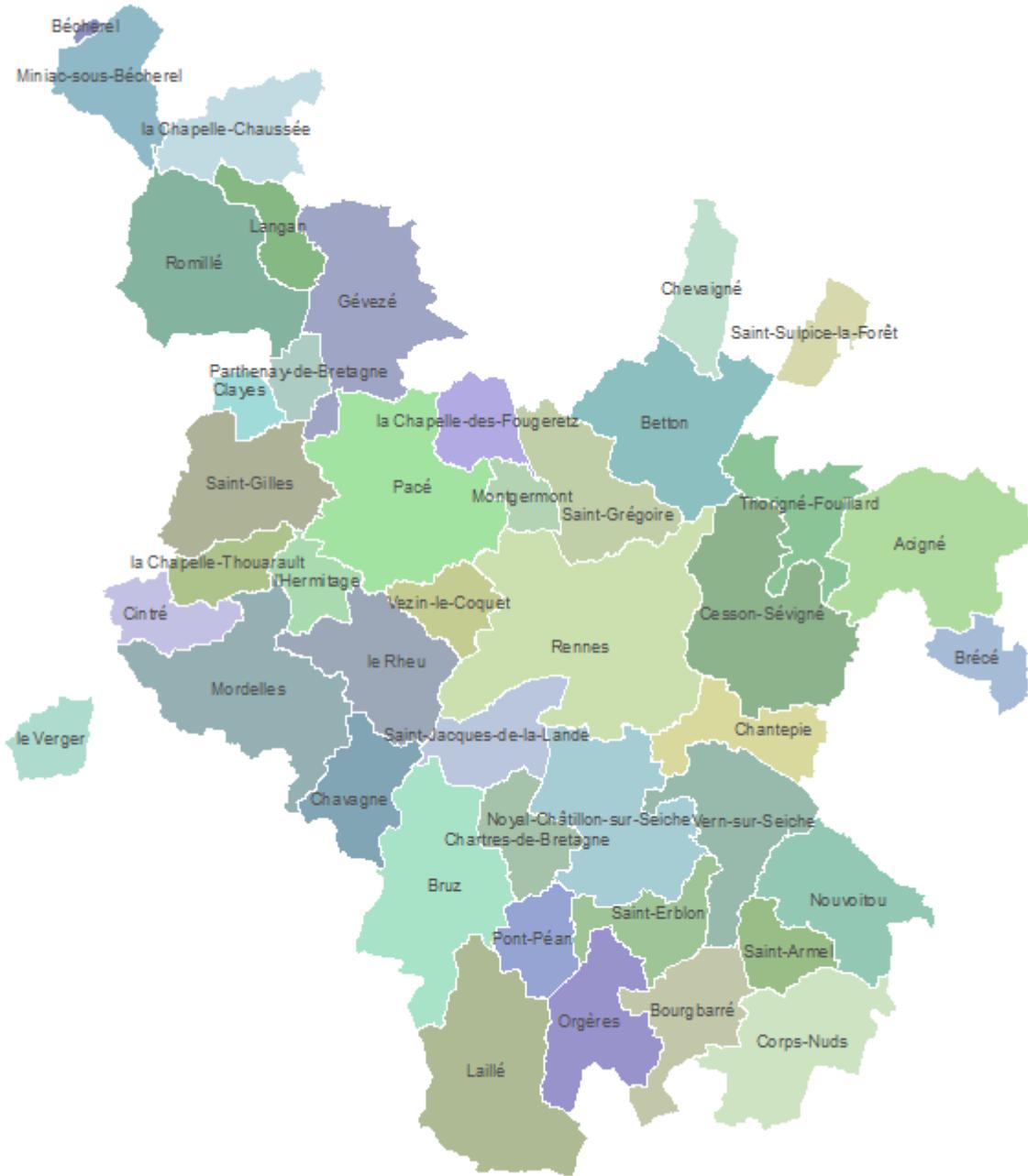


Figure 3 : Cartographie des communes de Rennes Métropole

I.1.1 Domaines d'intervention

Le passage en métropole fait que Rennes Métropole renforce ses anciennes compétences et en acquiert de nouvelles.

Rennes Métropole intervient sur les domaines suivants :

- Développement économique et emploi :
 - Aide à l'emploi au soutien des acteurs économiques
 - Soutien à l'enseignement supérieur et à la recherche

- Aménagement et gestion de toute nouvelle zone d'activité
- Pilotage des pôles de compétitivité pour favoriser l'innovation
- Participation au capital des sociétés de transfert de technologie (de la recherche vers l'industrie)
- Aménagement et urbanisme de l'espace communautaire :
 - Organisation de l'espace communautaire et de son articulation avec l'ensemble du pays de Rennes
 - Réalisation d'études pour des projets utiles à l'ensemble du territoire
 - Conseil aux communes sur leurs projets
 - À terme (2019-2020), aura en charge un plan local d'urbanisme intercommunal effectif sur l'ensemble des communes
- Déplacements :
 - Définition et suivi d'un plan pour les vélos
 - Aide au développement des voitures électriques
 - Organisation des transports en commun (bus et métro)
 - Participation à la gouvernance et à l'aménagement des gares
- Numérique :
 - Soutien à la recherche et aux entreprises innovantes du numérique
 - Pilotage de l'ensemble des infrastructures et de l'aménagement du réseau de télécommunication amenant le web et les services liés aux habitants, entreprises et laboratoires de recherche
- Habitat :
 - Adoption d'un programme local de l'habitat pour accueillir les nouveaux habitants et améliorer les logements
 - Mise en œuvre d'une politique équilibrée et cohérente en matière de logement social
 - Création et gestion des lieux d'accueil des gens du voyage
 - Amélioration du parc immobilier et réduction de l'habitat insalubre en lien avec les communes
- Environnement :
 - Prévention de la production de déchets, traitement des déchets (y compris suivi des déchèteries)
 - Protection de l'environnement (actions pour réduire la pollution, qualité de l'air, bruit...)
 - Rôle accentué dans la transition énergétique sur son territoire
- Culture :
 - Construction, gestion et entretien des équipements culturels d'intérêt communautaire (les Champs Libres, écomusée...)
 - Organisation et soutien des actions en faveur de la culture qui ont un intérêt pour le territoire

- Cohésion sociale :
 - Contrats de ville pour les habitants vivant dans les zones les plus en difficulté
 - Accès à la culture et au sport (carte 'Sortir !')
- Coopération et rayonnement :
 - Coopération avec Nantes, Saint-Malo, le Québec
 - Participation au Pôle métropolitain Loire-Bretagne
 - Construction, aménagement et gestion du centre de congrès
- Eau et assainissement :
 - Préservation de la qualité de l'eau et production d'eau potable
 - Traitement des eaux usées
- Gaz, chaleur, électricité :
 - Distribution du gaz et de l'électricité
 - Réseau de chaleur (chauffage urbain)
- Routes :
 - Création, entretien et aménagement des routes communales
 - Éclairage public
 - Chemins dédiés aux vélos (après 2017)

I.1.2 Instances et rôles¹

Le conseil communautaire : c'est le lieu du débat et des décisions les plus importantes de l'institution. Composé de 122 représentants des communes membres, il se réunit une fois par mois.

Le bureau métropolitain : pour alléger le travail des élus et accélérer les prises de décisions, le conseil lui a délégué la possibilité de prendre des décisions dans des domaines précis (mise en œuvre du programme d'action foncière par exemple). Organe de composition pluraliste, il joue également un rôle consultatif : il donne son avis sur les projets de délibérations soumises au conseil. Ses 27 membres (le président, 20 vice-présidents et 6 conseillers communautaires) se réunissent une fois par mois pour préparer le conseil.

Les commissions restreintes : elles sont composées de conseillers métropolitains et sont présidées par le vice-président en charge du domaine concerné. Elles participent à la préparation des décisions du bureau métropolitain et des délibérations du conseil communautaire.

Les commissions élargies : elles sont ouvertes aux élus des communes. Elles permettent la communication dans les communes concernant les projets métropolitains et les politiques publiques.

¹ Informations en partie récupérées à l'URL : <http://metropole.rennes.fr/politiques-publiques/elus-institution-citoyennete/institutions-et-competences/les-instances-metropolitaines/>

Les comités de secteur: au nombre de 8, ils sont composés d'élus de chaque commune. Il s'agit de lieux de débat vis-à-vis des politiques publiques afin de favoriser la coopération, la mutualisation entre les communes.

Les comités de pilotage et groupes de travail: mis en place par le président de la métropole. Une lettre de mission précise leurs objectifs sur des sujets transverses et spécifiques. Ils sont composés d'élus municipaux et communautaires mais peuvent également être composés de personnalités qualifiées.

La conférence des maires: est présidée par le président de RM. Elle est composée des 43 maires des communes, et se réunit une fois par trimestre. Il s'agit d'un lieu d'échange et de débat sur les axes métropolitains et les projets structurants.

La convention des élus: se réunit une fois par an, et est composée de l'ensemble des élus de la métropole. Elle est un lieu de débat sur les projets structurants et les orientations de la métropole.

I.2 La Ville de Rennes

I.2.1 Quelques chiffres

Capitale de la Bretagne, Rennes comptait plus de 213 956 habitants au 1^{er} janvier 2015 (source INSEE). Il s'agit de la onzième plus grande ville de France en nombre d'habitants. Rennes est la ville la plus peuplée de la région Bretagne devant Brest.

Rennes est une ville universitaire, près de 60 000 étudiants y viennent étudier chaque année.

Elle est le chef-lieu de Bretagne et d'Ille et Vilaine. À ce titre elle accueille le siège du Conseil Général d'Ille et Vilaine, du Conseil Régional, la préfecture.

Cette administration emploie près de 4000 personnes réparties dans 9 directions générales.

I.2.2 Instances et rôles

Le conseil municipal : c'est le lieu du débat et des décisions les plus importantes de l'institution. Il est composé de 61 représentants élus au suffrage universel.

Le maire : Exécutif du conseil et chef des services de l'institution, il prend également, par délégation du conseil, des décisions limitativement énumérées.

Depuis 2010, Rennes Métropole et la Ville de Rennes ont mutualisées plusieurs de leurs services, dont la Direction des Systèmes d'Information. Elles emploient à elles deux plus de 5000 agents.

I.3 Direction des Systèmes d'Information

La DSI Ville de Rennes a été créée en 1976 avec 4 personnes et avait pour principale mission à l'époque, l'informatisation de certains services municipaux de la Ville de Rennes.

L'informatisation du système d'Information de la Ville a commencé par la mise en place du système informatique comptable.

Parallèlement, la DSI de Rennes Métropole fut créée en 1996. Au départ le service était composé d'une seule personne.

Au fil du temps les missions de ces 2 services se sont étendues et leur effectif a augmenté. En juillet 2010, suite à une volonté politique, certains services de la Ville de Rennes Métropole et de la Ville, dont la DSI, ont été mutualisés. Ainsi, les équipes des 2 services informatiques ont fusionné. C'est ainsi qu'est née la Direction mutualisée des Systèmes d'Information.

Aujourd'hui ses missions sont les suivantes :

- Mettre à disposition des services les moyens matériels et logiciels (informatiques et télécoms) nécessaires à leur activité
- Concevoir l'architecture des systèmes d'informations en privilégiant la cohérence, l'ouverture vers les usagers et les partenaires tout en garantissant la sécurité
- Assurer l'exploitation et la maintenance des systèmes informatiques et télécoms

Associée à ces missions, elle assure également certaines missions sur le pilotage du Système d'information du Centre Communal d'Action Social (*CCAS*) qui dispose d'une équipe informatique dédiée.

Ci-dessous, quelques indicateurs chiffrés représentant la DSI mutualisée.

- Effectifs : 57 personnes
- Sites : 3 sites « cœur de réseau » raccordés via le réseau FOR (*Fibre Optique Rennaise*) en 10 Gb/s, 5 sites de distribution via le réseau FOR en 1 Gb/s, et enfin 126 sites distants raccordés soit au réseau FOR, soit en SDSL-VPN ou ADSL-VPN.
- Micro-ordinateurs : 2722 (Ville de Rennes) + 975 (Rennes Métropole)
- Imprimantes : 1072 (Ville de Rennes) + 61 (Rennes Métropole)
- Téléphones fixes : 3589 (Ville de Rennes et Rennes Métropole) + 423 (Les Champs Libres)
- Téléphones mobiles : 1160 (Ville de Rennes, dont 160 smartphones) + 155 (Rennes Métropole, dont 26 smartphones)

La DSI mutualisée est divisée en 4 services principaux et comprend un pôle basé dans l'équipement culturel : *Les Champs Libres*. Les 4 services sont subdivisés en équipes (*cf. Organigramme de la Direction des Systèmes d'Information*). Physiquement, la majorité du personnel est implanté au 38 rue Vasselot (excepté les agents des Champs Libres (4 personnes) et ceux du support de l'Hôtel de Rennes Métropole (2 agents)).

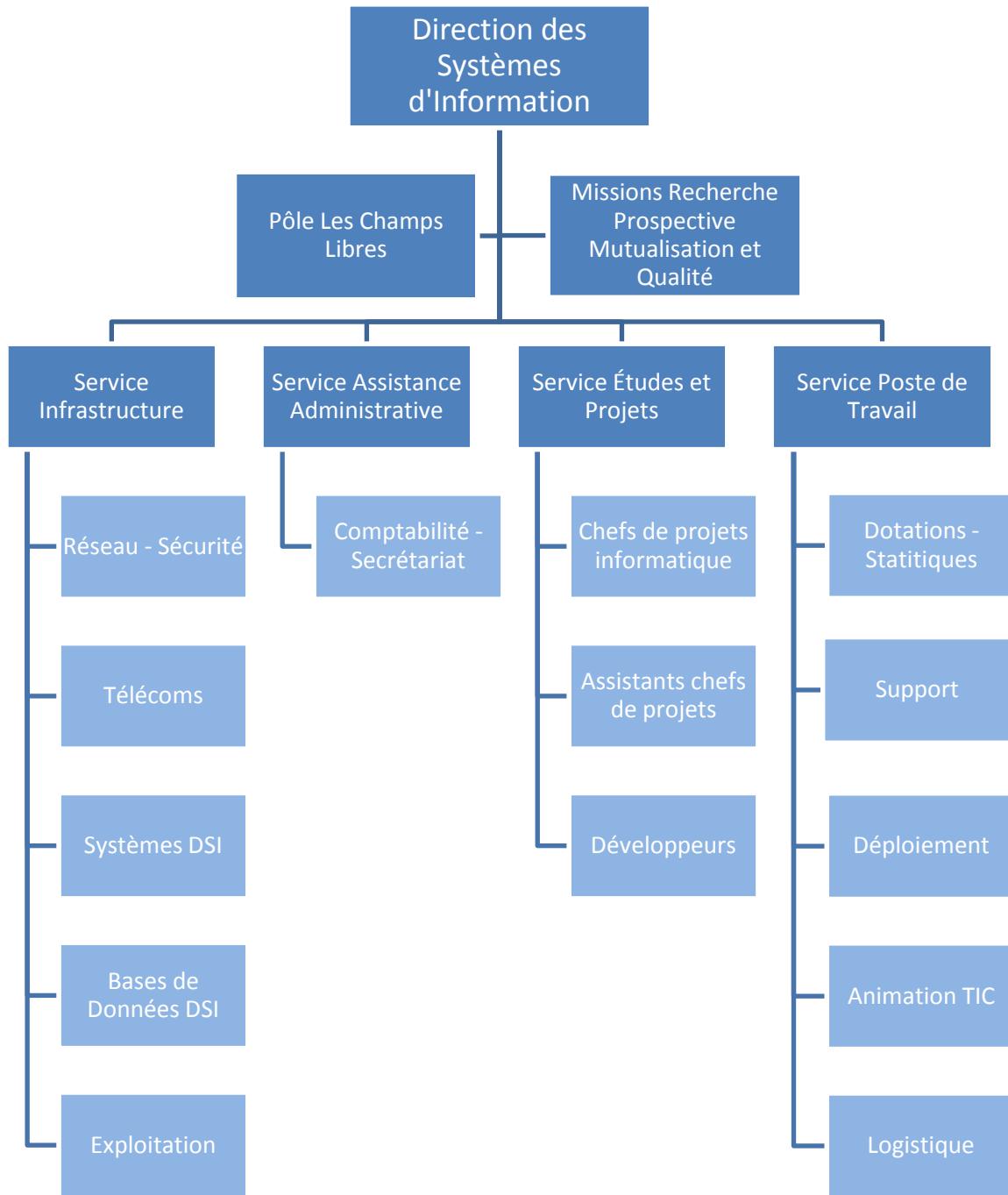


Figure 4 : Organigramme de la Direction des Systèmes d'Information

I.3.1 Le Service Etudes et Projets

Les missions du service Etudes et Projets peuvent être résumées par **la définition, la mise en œuvre et le maintien en condition opérationnelle de solutions logicielles métiers dans le cadre :**

- de ses bénéficiaires (*agents utilisateurs et leurs contraintes, usagers*)
- de projets du mandat politique (*élus et DG*)
- de la politique informatique définie par le DGS
- de ses opérateurs (*autres maîtrises d'œuvre, éditeurs, partenaires, autres services DSI*)
- des environnements concernés internes et externes (*cohérence et urbanisation du système d'information, environnements techniques, sécurité, environnements liés aux autres maîtrises d'œuvre*)
- de ses contraintes économiques (*budget et temps homme*)
- de ses contraintes juridiques

Le service Etudes et Projets compte aujourd’hui 19 agents.

Tous les agents sont rattachés hiérarchiquement au responsable de service mais travaillent les uns avec les autres, créant ainsi des « associations ponctuelles » en mode projet. Les agents sont plutôt spécialisés dans un ou plusieurs domaines fonctionnels (finances, ressources humaines, droit des sols...) mais peuvent ponctuellement intervenir en dehors de leur domaine de prédilection.

Cette organisation permet d'affecter, avec flexibilité, des ressources à un projet en fonction essentiellement des compétences techniques, des compétences fonctionnelles et de la disponibilité des agents.

On dénombre 3 grands profils dans le service:

- Les chefs de projets
- Les assistants de chefs de projets
- Les analystes développeurs

Les chefs de projets sont au nombre de 8 dans le service. Leur rôle consiste à mener de bout en bout la mise en œuvre et le suivi de logiciels spécialisés dans des métiers exercés dans les collectivités Ville de Rennes et Rennes Métropole. Pour cela, leurs activités se décomposent principalement de la manière suivante :

- La rédaction des cahiers des charges
- Le suivi des consultations et des marchés sur l’acquisition des logiciels
- Le suivi de la mise en œuvre opérationnelle des logiciels

- Le suivi des évolutions
- Le suivi de la résolution d'anomalies
- L'expertise sur leurs domaines fonctionnels

Les assistants de chefs de projets sont au nombre de 4. Ils ont des compétences techniques transversales telles que la connaissance du langage de requêtes SQL (*Structured Query Language*) pour les bases de données, la construction de statistiques ou de rapports. Ils peuvent gérer des petits projets sous la responsabilité d'un chef de projet ou du chef de service. Ils ont véritablement un profil technique.

Ils ont une connaissance fonctionnelle des logiciels qu'ils gèrent et peuvent donc effectuer du conseil en paramétrage, de l'administration fine, du dépannage, un contact de premier niveau avec les éditeurs, les tests et l'installation de nouvelles versions, le suivi des anomalies...

Ils constituent donc un premier niveau de réponse aux sollicitations des utilisateurs en évitant que le chef de projet ne soit sans cesse dérangé. Ils sont le relais du chef de projet dans les secteurs fonctionnels qu'ils connaissent. Faisant véritablement « équipe » avec celui-ci, ils contribuent à la sécurité des applications.

Les analystes développeurs, au nombre de 3, représentent l'équivalent de 2,5 temps plein réparti. Leurs rôles dépassent largement le cadre du développement d'applications : ils constituent en réalité le pool technique du service Etudes et Projets et sont largement sollicités par les chefs de projets et assistants/chefs de projets pour du conseil sur des causes de pannes ou lenteurs logicielles, des choix d'architecture technique, des requêtes pointues, des reprises de données... Ils sont également sollicités pour mener des expérimentations. Ils ne peuvent remplir ces missions que parce qu'ils pratiquent, expérimentent au quotidien dans le cadre des applications qu'ils produisent et maintiennent. Des développements leur sont confiés uniquement dans les cas suivants :

- Secteur délaissé par les éditeurs ou solutions non pertinentes
- Interface avec des logiciels existants
- Test de technologies

I.3.2 Le Service Infrastructure

Le service Infrastructure est composé de 17 agents répartis dans 5 équipes ayant chacune des missions spécifiques. Le point commun entre elles : l'infrastructure du Système d'Information de la Ville et de Rennes Métropole.

L'équipe réseau et sécurité, composée de 3 agents, est en charge de l'administration, la mise en œuvre et la maintenance du réseau fusionné de la Ville de Rennes et Rennes Métropole

au niveau matériel et paramétrage logiciel. Elle a également des missions de sécurité telles que la non-intrusion d'éléments non autorisés sur le réseau professionnel.

L'équipe système gère essentiellement les serveurs, leur sécurité, le stockage des données, l'infrastructure de messagerie, la tenue des annuaires et les droits utilisateurs sur le système de fichiers. 4 agents travaillent au sein de cette équipe

L'équipe télécoms gère la mise en œuvre et la maintenance du système téléphonique, les marchés, la facturation Télécom, les annuaires téléphoniques (pabx) ainsi que les câblages et les actifs réseaux. Elle assure également des missions de support sur les aspects téléphoniques. Enfin elle prend en charge les opérations informatiques liées à un déménagement ou un mouvement de personnel. 5 personnes composent cette équipe.

L'équipe base de données, composée de 3 agents affectés chacun à mi-temps sur ce sujet, a pour mission de sécuriser l'accès aux bases de données, d'optimiser leur fonctionnement, d'assurer diverses tâches d'administration comme la sauvegarde des données. L'équipe gère un parc important de bases de données hétérogènes (Oracle, SQLServer, PostgreSQL, MySQL).

L'équipe exploitation représente l'équivalent de 2 temps plein. Ses missions sont d'assurer l'ensemble des tâches d'exploitation (exécution des traitements batchs, sauvegardes...), la surveillance des différents serveurs à l'aide d'outils de supervision. Elle gère également les impressions de masse (factures, bulletins de paie, listes électorales...). Elle s'occupe également des droits d'accès des utilisateurs au référentiel Business Object XI.

I.3.3 Le Service Poste de Travail

Le service Poste de Travail est composé de 14 personnes. L'ensemble des missions exercées par les différentes équipes ont toutes un rapport avec le poste de travail de l'utilisateur :

L'équipe Dotations - Statistiques gère les marchés relatifs au service Poste de Travail pour l'achat de matériel informatique et de logiciels non métier. Elle s'occupe également des achats liés à ces marchés ainsi que de leurs facturations. Parallèlement elle prépare les plannings de déploiement des postes utilisateurs à renouveler pour l'équipe Déploiement. Cette équipe représente l'équivalent d'un temps plein.

L'équipe Déploiement se compose de 3 personnes. Elle s'occupe de la préparation des nouveaux postes utilisateurs (installation du Système d'exploitation, des logiciels bureautiques, des logiciels métiers) dans le cadre des renouvellements de postes. Elle a en charge la conception du déploiement de scripts ou packages, du déploiement d'applicatifs sur l'ensemble du parc. Enfin elle coordonne le travail du sous-traitant en charge de l'installation physique des postes dans les services. Les déploiements « complexes » sont en effet gérés en interne alors que les autres sont externalisés.

L'équipe Animation TIC, composée de 2 agents travaillant à mi-temps sur ces missions, s'occupe d'organiser et d'animer des ateliers bureautiques sur différents logiciels informatiques comme la suite bureautique Office par exemple. Elle conseille les services en matière d'organisation de répertoire, d'utilisation de la bureautique et gère également une bibliothèque de documents de référence mise à disposition de tous les utilisateurs d'outils informatiques de la Ville de Rennes.

L'équipe Support, assure les fonctions habituelles de support auprès des utilisateurs de postes informatiques et de smartphones. 7 agents y travaillent dont 2 sont localisés à l'Hôtel de Rennes Métropole.

L'équipe Logistique est une équipe composée de 2 agents. Ses missions principales sont de saisir des informations relatives au poste de travail de l'utilisateur dans le logiciel de Gestion de Parc de la DSI (logiciels installés...), d'analyser les besoins relatifs à certaines demandes utilisateurs ou à des mouvements de personnel. Elle s'occupe également de rationaliser les moyens d'impression dans les services. Elle gère aussi les prêts de matériels tels que les portables ou smartphones. Enfin elle effectue une mission de sécurité, qui consiste à relever les alertes remontées par l'antivirus centralisé et de prendre les décisions adéquates (remplacement du poste, contact avec l'utilisateur concerné...).

I.3.4 Le Service Assistance Administrative

Il constitue la partie administrative de cette direction très technique. Il se compose d'un seul service : le service « Comptabilité - Secrétariat ». Il assure une assistance dans la mise en place des marchés, effectue le suivi des factures et de la comptabilité du service, et réalise les missions d'accueil. 4 personnes y travaillent.

Concernant mon affectation, je travaille à mi-temps en tant qu'analyste-développeur au sein du service Études et Projets, et en tant qu'administrateur de bases de données du service Infrastructure sur l'autre mi-temps.

I.4 Service des Archives Municipales

Le service des archives municipales fait partie de la direction générale de la culture (DG Cult). Ce n'est pas un service mutualisé, il est donc uniquement Ville de Rennes. Toutefois, par convention de mise à disposition de services, il exerce également des missions pour Rennes Métropole. Il est composé de 16 agents.

C'est un service patrimonial dont les missions sont de collecter, conserver et mettre à disposition du public les dossiers constitués par les services de la Ville de Rennes et les offices para-municipaux. Il intègre également désormais le fonds d'archives de Rennes Métropole.

Enfin, les archives municipales conservent un important fonds ancien, constitué des archives de la communauté de ville, dont les premiers documents remontent au début du XVe siècle. Contrairement à d'autres collectivités de tailles plus réduites, les archives définitives sont en effet conservées par les archives municipales et non par les archives départementales.

Ces archives doivent satisfaire un public multiple : administration, chercheurs, historiens et un public plus large, constitué de non-spécialistes, pour lequel le service doit jouer un rôle de médiateur entre le document brut et le discours historique par le biais d'actions diverses (expositions, dossiers pédagogiques, publications, conférences...).

Le service des archives élabore en collaboration avec les autres services le tableau de gestion. Il s'agit d'un document contractuel qui recense tous les documents produits par un service, en les organisant par grande mission. Pour chaque document, il fixe la durée d'utilité administrative, c'est à dire la durée minimum pendant laquelle le document doit être conservé. Il précise également le sort final du document, la destruction ou le versement aux Archives de Rennes pour conservation définitive.

Ces deux points sont déterminés en fonction de la législation en vigueur, des délais de prescription, et de l'expertise de l'archiviste et des agents concernés. Il est donc nécessaire de mettre régulièrement à jour le tableau de gestion.

I.5 Mégalis Bretagne

Mégalis est un syndicat mixte ouvert créé en 1999. Il est composé de 107 membres. Le Conseil Régional de Bretagne, les Conseils Généraux des Côtes d'Armor, du Finistère, d'Ille-et-Vilaine et du Morbihan, 2 métropoles (Rennes Métropole et Brest Métropole), 9 communautés d'agglomération et 92 communautés de communes.

Conformément à ses statuts, ses compétences sont :

- animer et gérer le projet Très Haut Débit Bretagne
- encourager le développement des usages des réseaux de communication électronique et favoriser le développement de l'administration électronique

Nous venons donc de présenter les collectivités pour lesquelles je travaille, ainsi que les services et autres interlocuteurs qui interviennent dans le projet d'archivage électronique.

II État de l'art

Le développement constant des nouvelles technologies a fait du numérique un outil omniprésent dans tous les domaines. Pour donner un exemple, la photo argentique a totalement été remplacée par son « reflet » dans le monde numérique. La dématérialisation des processus des collectivités territoriales en est également une preuve.

Cette transition numérique facilite grandement l'accès à l'information, rendant les processus plus fluides. Toutefois, de nouvelles vulnérabilités et problématiques surgissent pour la conservation des documents numériques tout au long de leur cycle de vie.

L'objectif de cette partie est de dresser un état de l'art en matière d'archivage numérique à l'heure actuelle via un panorama des diverses problématiques rencontrées et les solutions proposées pour y faire face, à travers des contraintes juridiques, techniques ou normatives, afin de nous permettre par la suite de définir des spécifications fonctionnelles et techniques, et par conséquent de choisir une solution adaptée pour la maquette à réaliser.

Avant d'aborder la préservation de l'information numérique face à l'obsolescence technologique, à travers les supports de stockage ou les formats de fichiers préconisés, nous effectuerons une succincte présentation des archives, ainsi qu'une présentation du cadre juridique et normatif.

Il faut néanmoins prendre conscience que le domaine de l'archivage électronique est un secteur en pleine mutation de fait de son jeune âge. Par conséquent, cet état de l'art se veut valable à l'heure où sont rédigées ces lignes mais est susceptible d'évoluer dans les prochains mois, notamment dans le domaine juridique traitant de l'archivage patrimonial.

II.1 Présentation des archives

Une archive désigne un original, et non une copie, qui sera conservée sur le moyen ou long terme selon les durées de conservation définies par le code du patrimoine.

Dans le cadre de l'archivage électronique, une archive peut-être un fichier, une donnée, un document numérisé ou un ensemble de documents.

En France il existe des archives datant de l'antiquité ou du moyen-âge sous forme de tablette d'argile ou encore de papyrus. Mais c'est quelques années après la révolution française qu'a été créée une véritable institution pour gérer ces archives, les archives

nationales. La Direction des Archives est apparue en 1897 et est devenue la Direction des Archives de France en 1936 [4]¹.

II.1.1 Finalités des archives

Les archives ont plusieurs finalités. Tout d'abord, l'intérêt d'une archive peut être son intérêt historique, patrimonial, statistique ou encore scientifique. Elle sert alors à la recherche, à l'histoire de la commune, du département, du pays.

Les archives contiennent en effet des preuves du passé, de l'histoire. Elles constituent la mémoire d'un pays, d'une nation, de l'humanité.

À titre d'exemple, la carte numérisée ci-dessous nous montre la carte de Rennes en 1616, pouvant être forte intéressante pour des chercheurs ou historiens :



Figure 5 : Rennes, ville capitale de Bretagne et siège du Parlement 1616²

Les archives peuvent également servir de preuve dans un litige. On parle alors d'utilité administrative de l'archive. Par exemple un contrat, une facture ou un récépissé peut faire valeur de preuve devant un juge. Ces documents sont généralement signés et ont été envoyés ou reçus par quelqu'un à une certaine date.

¹ Direction des Archives : Historique des archives

² Image récupérée sur le site des Archives de Rennes, cote 1FI42 (<http://www.archives.rennes.fr>)

II.1.2 Actions sur une archive

Plusieurs actions sont possibles sur une archive électronique. Pour commencer, une archive est versée dans un système d'archivage, après évaluation de sa conformité (à savoir est-ce qu'elle mérite bien d'être archivée et correspond-elle aux critères prédéfinis par le service compétent?).

Ensuite, l'archive est triée et classée selon un plan de classement préétabli. Elle peut ensuite être consultée, des outils de recherche permettant de la trouver.

Enfin, l'archive subit son sort final. Elle est alors éliminée ou conservée si elle est jugée représenter un intérêt historique ou patrimonial.

Ces actions interviennent à des stades différents de la vie de l'archive, on parle du cycle de vie de l'archive. Dans le domaine de l'archivage, on parle de la théorie des 3 âges, qui court de la création de l'archive jusqu'à son sort final.

II.1.3 Théorie des 3 âges

La théorie des 3 âges, d'origine anglo-saxonne, a été transposée dans le monde archivistique français en 1948 par Yves Pérotin, archiviste et théoricien français. Elle traite du cycle de vie de l'archive sous les 3 âges, qui sont les **archives courantes**, les **archives intermédiaires** et les **archives définitives**.

Certains acteurs de l'archivage électronique indiquent qu'il n'est plus pertinent de parler de cette théorie pour l'archivage électronique [5]¹.

Ils postulent qu'un document électronique à archiver doit en effet être pris en compte dès sa création afin de pouvoir évaluer son intégrité, et que par conséquent la notion d'archives intermédiaire n'a plus beaucoup d'intérêt.

Cette théorie n'est pas à rejeter mais elle doit plutôt être transposée sur les responsabilités vis-à-vis du document pendant son cycle de vie.

II.1.3.1 Archives courantes

L'âge des archives courantes correspond au moment à partir duquel est créé le document. Cette période court tant que le document est fréquemment accédé, et qu'il reste donc utile quotidiennement. On parle de durée d'utilité courante (*DUC*).

Dans le cadre de ce projet, cette période correspondra à la durée pendant laquelle le document fait partie de la **GED**.

¹ Papy, F. : *Recherches ouvertes sur le numérique*

Le ou les producteurs, ainsi que d'autres acteurs sont susceptibles de le modifier et d'y accéder.

II.1.3.2 Archives intermédiaires

Lorsque ces documents cessent d'être utiles quotidiennement, et surtout qu'ils ne nécessitent plus de modification, ils deviennent alors des archives intermédiaires. Ils ne sont plus modifiables mais peuvent toutefois être accéder ponctuellement pour diverses raisons.

Cette période recouvre notamment la notion d'archivage probant ou archivage légal, période pendant laquelle la conservation des documents est réglementée et obligatoire. Elle est donc régie par la DUA du document, et ce à des fins administratives, fiscales ou encore d'information.

Afin de conserver une notion de preuve, de fournir des garanties techniques et juridiques ayant les mêmes valeurs qu'un document papier, ce type d'archive nécessite une sécurisation des documents.

Pour prouver la valeur d'un document, il faut avoir mis en place différentes pratiques suivantes :

- l'authenticité et l'engagement des parties à l'aide de certificat électronique (ou signature électronique), d'horodatage et de calcul d'empreintes.
- l'intégrité du document à l'aide d'empreinte de hachage afin de prouver qu'il n'a pas été modifié.
- l'horodatage du document afin de prouver sa provenance et de conserver sa non-répudiation.
- la traçabilité afin de connaître les mouvements qui ont eu lieu sur le document, qui est également traitée avec des calculs d'empreintes basées sur des algorithmes de hachage (comme MD5¹ ou SHA-2²).
- la confidentialité via du cryptage et du contrôle de droits d'accès, des mécanismes d'authentification.
- la lisibilité via des mécanismes de migration des formats et des supports afin de rendre l'archive accessible et lisible durant toute sa vie. Cet aspect implique également la compréhensibilité de l'archive, qui peut être garantie par ce que l'on nomme les métadonnées. Une métadonnée est une donnée décrivant les données. Elles peuvent décrire aussi bien le contexte du document, son intérêt, son créateur, son format numérique...

¹ Message Digest

² Secure Hash Algorithm

- l'accessibilité via des méthodes de sauvegarde, de réPLICATION permettant de s'assurer de la conservation du document.

En plus de ces qualités essentiellement liées à la conservation du critère de preuve, l'archive doit également être réversible. C'est-à-dire qu'il doit être possible de la récupérer, pour l'intégrer par exemple dans un SAE définitif ou encore pour effectuer une migration vers un autre système.

II.1.3.3 Archives définitives

Au terme de la DUA le document subit son sort-final. S'il n'a plus de valeur juridique, patrimoniale, scientifique ou statistique, son sort-final sera l'élimination. Sinon, il devient une archive définitive et est conservé indéfiniment.

Ces archives ont donc déjà subi un traitement particulier. Elles sont passées par des procédures de tri en amont afin de vérifier qu'elles devaient bien être conservées.

Un document peut également devenir définitif dès sa création, cela peut être le cas par exemple pour les documents d'état civil. Il ne passera pas dans ce cas par l'archivage intermédiaire, mais sera directement intégré à un SAE patrimonial.

Le schéma ci-dessous résume ce cycle de vie des 3 âges :

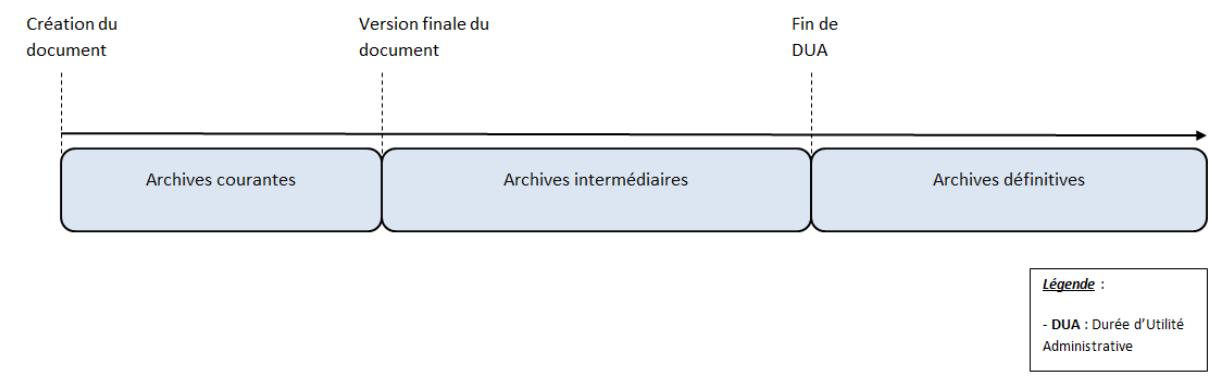


Figure 6 : Théorie des 3 âges - cycle de vie du document

II.1.4 Record Management

Contrairement au français, en langue anglaise le terme « archive » ne débute qu'à l'âge des archives définitives. L'âge courant et intermédiaire du document est couvert par le terme record management (RM). Ce terme est important car il est beaucoup utilisé, même en langue française, et beaucoup de normes traitent du record management.

Le RM ne correspond donc pas à l'archivage définitif mais peut être interprété (en France) par la gestion des documents d'archives. Il couvre l'organisation des documents de leur création à leur réception, leur utilisation et leur sort final (destruction ou transfert vers le définitif). On parle également d'archivage managérial. Il traite par conséquent la prise en compte en amont des processus d'archivage dans le but de satisfaire la traçabilité, l'intégrité, la sécurité ou encore la pérennité des documents.

II.2 Textes de référence

Dans ce chapitre nous allons présenter les principaux textes de référence couvrant le domaine de l'archivage électronique.

Nous effectuerons ce panorama en parcourant le cadre juridique, ainsi que les normes et standards à respecter dans le cadre de l'archivage électronique.

L'océan normatif en matière d'archivage et d'archivage électronique est très vaste, et il est facile de s'y noyer. Par conséquent, nous allons présenter les principales normes du domaine de l'archivage électronique, et nous tenterons de montrer les liens entre elles et leur place dans la mise en place d'un SAE.

Il faut toutefois préciser qu'une partie de ces normes étant payantes, nous ne pourrons pas en montrer des exemples précis, mais seulement en ressortir les résumés qui sont proposés par les organismes éditeurs.

II.2.1 Cadre juridique et réglementaire

Le cadre juridique de l'archivage repose essentiellement sur le code du patrimoine (partie législative et partie réglementaire). Nous ne comptons pas présenter ici tous les textes de lois ou décrets, traitant de près ou de loin des archives.

Nous présentons essentiellement les principaux textes ayant un impact fort sur le choix d'une solution lors de la mise en place d'un SAE dans une collectivité territoriale. Nous invitons par conséquent le lecteur désireux d'aller plus loin de se référer au code du patrimoine et au code civil.

Les modalités d'archivage de documents électroniques ne sont régies par aucune loi actuellement. Toutefois, le Code Civil stipule que la fiabilité de l'archive est une condition obligatoire pour rendre le document valide, sans aucune précision sur la notion de fiabilité. Il convient alors de se baser sur le paysage normatif existant afin de mettre en œuvre les bonnes pratiques et processus pour mener à bien un archivage électronique probant.

II.2.1.1 Notions liées à l'intégrité, la fiabilité, la sécurité

D'un point de vue juridique, l'archive est liée à la question de preuve, elle permet la justification des droits d'un acte quelconque jusqu'à expiration de son délai de prescription.

Selon le code civil, « *l'écrit sous forme électronique est admis en preuve au même titre que l'écrit sur support papier, sous réserve que puisse être dûment identifiée la personne dont il émane et qu'il soit établi et conservé dans des conditions de nature à en garantir l'intégrité »* (**C.civ, art.1316-1**). L'identification de l'auteur (ou du destinataire) doit passer par un procédé fiable d'identification (**C.civ, art.1316-4**), soit la signature électronique du document, la certification de ce procédé étant confié à des tiers de confiance et/ou des tiers d'horodatage (**D., 30 mars 2001, n°2001-272**).

Il faut toutefois distinguer un document numérique et un document numérisé. Pour le premier, l'original est sous format électronique, alors qu'un document numérisé est une copie d'un document original sous forme papier.

Pour avoir la même valeur qu'un document papier, un document numérisé doit être la copie conforme de l'original. Il doit également être signé électroniquement, tout comme un document numérique par ailleurs. En effet, il faut pouvoir prouver l'identité de son auteur.



Figure 7 : Image humoristique sur la numérisation des archives¹

Le document numérique doit également être horodaté à sa création, et la copie numérique doit l'être le plus rapidement possible après sa numérisation (**Cass., Civ 2, 12.04.2008, n°07-17622, Bull. 2008, II, n°259**).

¹ Image récupérée sur le site <https://bouzou.wordpress.com/2010/08/02/humour-archivage/>

Les délais de communicabilité des archives publiques sont aussi fixés par la loi (**L. n° 2008-696, 15 juil. 2008**). Avant ce délai, des droits d'accès aux documents d'archives doivent par conséquent être mis en place, afin de restreindre lesdits accès.

II.2.1.2 Hébergement des archives électroniques

Depuis 2008, les archives courantes et intermédiaires peuvent être conservées par des prestataires privés (**L. n° 2008-696, 15 juil. 2008**). Le tiers archiveur doit se voir accorder l'agrément nécessaire à la conservation des archives par le SIAF. Cet agrément est délivré si le prestataire répond aux exigences normatives relatives aux prestations d'archivage et, s'il s'engage à conserver les archives sur le territoire national, ainsi qu'à mettre en place une politique de sécurité, de conservation et de confidentialité des archives déposées (norme NF Z 42-013).

Par contre, les archives définitives doivent à l'heure actuelle être obligatoirement conservées par la collectivité ou le service public d'archives compétent dont elles dépendent, généralement les archives départementales. Dans notre cas, les Archives Municipales de Rennes gèrent eux-mêmes les archives patrimoniales.

Il est tout de même à noter que l'AAF (*Association des Archivistes Français*) essaye de faire intégrer au projet de loi CAP (*Création, Architecture et Patrimoine*) un volet concernant les archives [6]¹. Parmi les points traitant des archives, elle souhaite des amendements pour offrir la possibilité de mutualiser, et donc externaliser, les archives publiques définitives.

Les données à archiver peuvent également concerter des données à caractère personnel, comme par exemple les documents d'état civil. Elles ne peuvent être conservées à long terme que si elles présentent un intérêt scientifique, statistique ou historique (**L. n°78-17, 6 janv. 1978, art. 6 et C. patr., art. L212-3**).

Enfin précisons que certaines règles concernant entre autre l'archivage électronique sont précisées par trois référentiels.

Tout d'abord par le Référentiel Général d'Interopérabilité (RGI) [6]², qui fixe les règles techniques permettant d'assurer l'interopérabilité entre les autorités administratives. Il détermine notamment les normes et les standards qui doivent être utilisés (*ord., 08*

¹ Girard, H. : *Projet de loi Crédit : ce que les archivistes veulent obtenir* (site *La Gazette des Communes*)

² Secrétariat Général pour la Modernisation de l'Action Publique : RGI - interopérabilité

déc.2005, n°2005-1516 et D., 2 mars 2007, n°2007-284), ainsi que les formats des fichiers archivés.

Ensuite par le Référentiel Général de Sécurité (RGS) [7]¹ qui fixe en partie les règles de sécurité pour la mise en place de certificats électroniques, d'empreintes, d'authentification ou encore d'horodatage (D., 02 fév. 2010, n° 2010-112).

Enfin, surtout dans le domaine public, tout système d'information et par conséquent les archives électroniques également, doivent répondre au Référentiel Général d'Accessibilité pour les Administrations (RGAA) [8]². Ce référentiel fixe les modalités pour rendre accessible à tous les documents communicables en ligne (L. n° 2005-102, 11 fév. 2005), et par déduction logique les documents d'un SAE.

En plus de ces trois référentiels, il existe tout un ensemble de normes couvrant l'archivage.

II.2.2 Cadre normatif et standards

Nous présenterons ici les principales normes pour la mise en œuvre d'un SAE, intermédiaire ou patrimonial. Concernant les normes couvrant les formats de documents ou les supports de stockage, nous ne les aborderons pas dans ce chapitre. Nous préciserons en effet les préconisations en la matière dans le chapitre suivant.

Avant de présenter ces normes, il peut être utile de préciser la distinction entre une norme et un standard dans la langue française. Contrairement à la langue anglaise dans laquelle les deux termes sont traduits par le même mot, *standard*, la langue de Molière différencie ces deux notions.

- Une norme est « *un document qui définit des exigences, des spécifications, des lignes directrices ou des caractéristiques à utiliser systématiquement pour assurer l'aptitude à l'emploi des matériaux, produits, processus et services* » [9]³. Elle est approuvée par des instances officielles, dont les plus connues sont l'Organisation Internationale de Normalisation (ISO) et l'AFNOR.

- Un standard diffère d'une norme dans le sens qu'il peut être une norme ou pas, il peut également le devenir. Il est défini par des groupes qui n'ont pas de mandats officiels

¹ Agence Nationale de Sécurité des Systèmes d'Information : Le Référentiel général de sécurité (RGS)

² Secrétariat Général pour la Modernisation de l'Action Publique : RGAA - Accessibilité

³ International Organization for Standardization : Normes

pour créer une norme. Il s'agit d'un ensemble de préconisations, de recommandations préconisées qui s'imposent parfois par l'usage. Parmi les organismes les plus notoires à l'origine de standards, nous pouvons citer l'Object Management Group (**OMG**), le World Wide Web Consortium (**W3C**), le Consortium Unicode ou encore l'Organization for the Advancement of Structured Information Standards (**OASIS**).

II.2.2.1 Normes managériales du record management

Comme nous l'avons déjà précisé précédemment, le record management traite de la mise en place d'une politique d'archivage, des processus d'archivage... Certaines normes, plutôt orientées management, définissent des cadres pour cette démarche.

Pour mettre en place une approche orientée processus, il peut être utile de se baser sur les normes **ISO 9001**, **ISO 9004** ou encore **ISO 10006** qui traite du management de la qualité, ainsi que des normes **ISO/CEI 27001** et **ISO/CEI 27002** traitant de la sécurité de l'information.

Également orientée processus, l'**ISO 30300** (et suivantes) permet de mettre en œuvre un système de gestion des documents d'activité (**SGDA**). Elle permet de définir les typologies de documents produits et gérés par un organisme.

Les documents constituent des actifs importants pour l'activité, qu'ils soient juridiques, contractuels, informationnels, stratégiques. La mise en place de cette norme permet la traçabilité des processus et documents qui en découlent.

II.2.2.1.1 ISO 15489

Publiée en 2001, la norme **ISO 15489** (Information et documentation « Record Management ») permet l'analyse du cycle de vie du document, de sa stabilisation jusqu'à son archivage. C'est la norme pivot, la norme de référence en matière d'archivage.

Elle traite du record management, et couvre autant l'archivage numérique que l'archivage papier. Elle permet avant tout d'assurer une politique de qualité permettant la traçabilité des documents. Elle fixe la démarche d'archivage et les outils à utiliser pour la mettre en œuvre. Elle est par conséquent orientée essentiellement management ou organisation, et ne traite pas des aspects techniques.

Elle est découpée en deux parties. La première traite des principes directeurs pour la politique d'archivage de l'information, la deuxième propose une démarche pour mettre en place le record management scindée en huit étapes.

L'**ISO 15489** ne couvre pas l'archivage définitif, elle ne traite que de l'archivage courant et intermédiaire.

La norme pivot **ISO 15489** est complétée par deux normes enfants.

Il s'agit d'une part de la norme **ISO 26122**, qui définit deux types d'analyses des processus d'archivage, une fonctionnelle et une séquentielle.

D'autre part, la norme fille **ISO 23081** qui traite également du record management. Cette dernière définit les métadonnées qui doivent être spécifiées par l'organisme pour l'archivage à valeur probante des documents numériques.

II.2.2.2 Normes de modélisation pour l'archivage électronique

II.2.2.2.1 OAIS (ISO 14721)

Le modèle **OAIS** (*Open Archival Information System ou Système Ouvert d'Archivage d'Information*) est un modèle de référence open source créé par le CCSDS (*Consultative Committee for Space Data Systems*), qui a donné lieu à la norme **ISO 14721** en 2003 (version payante de l'**OAIS** pour pouvoir obtenir une certification).

Il est devenu la norme à utiliser pour modéliser l'organisation d'un SAE. Il est d'ailleurs reconnu et utilisé dans le monde entier. En France par exemple, il est utilisé à la BnF (*Banque Nationale de France*), à la DAF (*Direction des Archives de France*), au CNES (*Centre National d'Études Spatiales*)... Il traite aussi bien de l'archivage intermédiaire que de l'archivage patrimonial.

Nous aurions pu faire un chapitre entier traitant du modèle OAIS, du fait de son importance dans la mise en place d'un SAE, ainsi que de sa densité. Nous en faisons ici un résumé et laissons le lecteur désireux de l'étudier plus en profondeur se référer à sa documentation complète¹.

Le modèle OAIS est un modèle lié à l'organisation à respecter afin de mettre en place un SAE patrimonial. Il ne traite pas de la manière de mise en œuvre mais est plutôt un outil

¹ Version 1 d'OAIS traduite en français à l'adresse suivante :
[http://public.ccsds.org/publications/archive/650x0b1s\(F\).pdf](http://public.ccsds.org/publications/archive/650x0b1s(F).pdf)

indispensable à sa réalisation. Il définit en effet des concepts et précise le vocabulaire, les rôles, les fonctions et les responsabilités au sein des processus d'archivage du SAE. Le RGI (*Référentiel Général d'Interopérabilité*) recommande d'ailleurs son utilisation dans le cadre de cette mise en place.

Dans le modèle OAIS, l'archive est positionnée au cœur de l'environnement du SAE, et peut être interfacée avec les trois acteurs suivants [14]¹ :

- le producteur entité qui produit le document à archiver, le transmet au SAE avec les bonnes informations.
- l'utilisateur qui peut effectuer des recherches et visionner des documents d'archive.
- le management (aussi appelé direction ou administration) qui définit les axes de politique d'archivage, l'arbitrage...

Afin de pouvoir se faire comprendre des différents métiers intervenant dans le processus d'archivage, le modèle OAIS est décomposé en deux modèles basés sur le langage UML (*Unified Modeling Language*), un modèle d'information et un modèle fonctionnel.

II.2.2.2.1.1 Modèle d'information

Afin d'être lisible sur le long terme, OAIS définit qu'une information numérique (une suite d'octets) doit être associée avec des informations de représentation, permettant entre autres de décrire le logiciel utilisé, le format du fichier ou encore le codage utilisé pour pouvoir lire la série d'octets. En effet un fichier dont les caractères sont encodés en ASCII, UTF8 ou en ISO 8859-1 ne sera lisible que si on connaît le bon encodage utilisé. Ces informations de représentation permettront de pouvoir restituer et comprendre le document numérique.

L'ensemble que forme l'information (les données contenues dans l'objet numérique) et ses informations de représentation est appelé le contenu d'information.

Par contre, les informations de représentation ne suffisent pas pour remplir les différentes missions que doit offrir une archive. Il faut en effet lui ajouter des informations de pérennisation. Le modèle OAIS les découpe en cinq catégories :

- les informations de contexte
- les informations de provenance
- les informations d'identification
- les informations de droits d'accès

¹ Banat-Berger F. & Huc C. (PIAF) : Section 5 : Modèle de référence OAIS et normes associées

- les informations d'intégrité

L'ensemble alors obtenu (l'objet numérique à archiver, les informations de représentation et les informations de pérennisation) forme un paquet d'information. Le modèle OAIS distingue trois types de paquets d'information [12]¹ :

- SIP : **Submission Information Package** paquet d'information à verser au SAE.
- AIP : **Archival Information Package** ou paquet d'information archivé. Il est à noter qu'un AIP peut être généré à partir de plusieurs SIP.
- DIP : **Dissemination Information Package** ou paquet d'information à diffuser.

Un paquet d'information à archiver (AIP) peut être représenté par le modèle suivant :

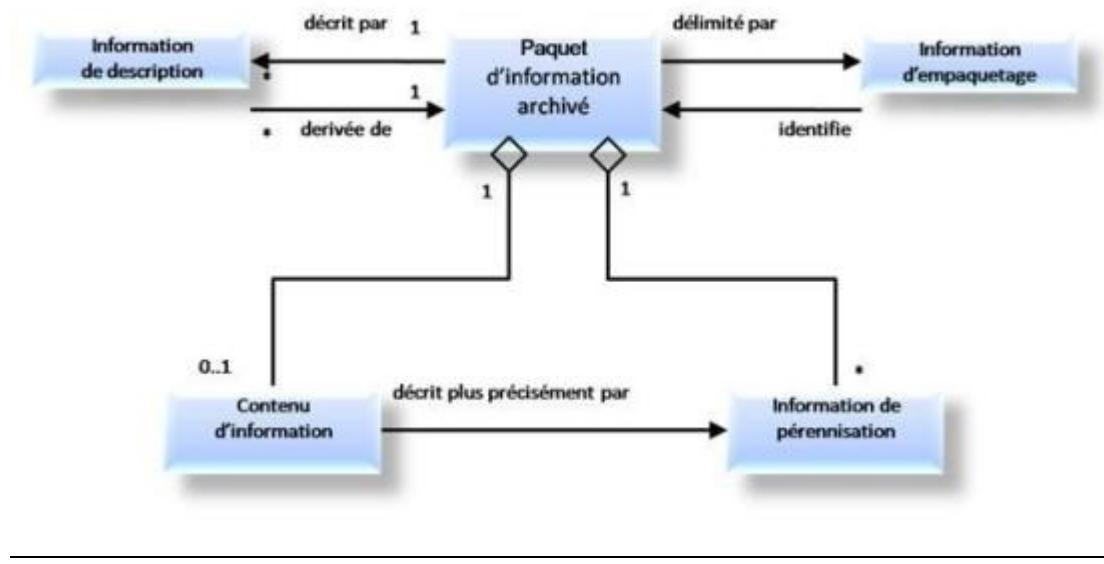


Figure 8 : Modèle du paquet d'information archivé (AIP)²

II.2.2.1.1 Modèle fonctionnel

Le modèle OAIS distingue plusieurs entités fonctionnelles :

- l'entité « Entrée » : qui traite de la réception des documents à archiver, de la vérification de leur conformité (format, informations de représentation et de pérennisation). Elle est en lien avec les entités « Gestion de données » et « Stockage ».

¹ Consultative Committee for Space Data Systems (CCSDS) : Modèle de référence pour un Système d'archivage ouvert (OAIS)

² Image récupérée sur le site du PIAF à l'adresse : http://www.piaf-archives.org/espace-formation/file.php/9/section05_web/co/module_section05_3.html

- l'entité « **Stockage** » : elle gère également la **conservation des données archivées** mais du **point de vue physique**. Elle traite des moyens de sauvegarde, de la migration des supports de stockage en cas de panne ou d'obsolescence technologique.
- l'entité « **Gestion de données** » : elle traite ce qui englobe la **conservation des données archivées** d'un **point de vue logique**, à savoir l'administration de la base de données, ses mises à jour, les fonctionnalités de recherche de documents ou encore les fonctionnalités de gestion des données à archiver.
- l'entité « **Administration** » : assure les **règles** définies au niveau du **SAE**. Elle permet la bonne **exploitation** du logiciel, les **migrations de données**.
- l'entité « **Planification de la pérennisation** » : permet de gérer les **recommandations de migration de support ou de format avant l'obsolescence technologique**. Elle gère également la modélisation des paquets d'information AIP et SIP.
- l'entité « **Accès** » : offre à un utilisateur **l'accès** à un document d'archive tout en vérifiant les **autorisations** d'y accéder. Elle gère aussi les résultats affichés à l'utilisateur lorsqu'il effectue une requête, comme par exemple savoir si un document précis est présent ou non dans le SAE.

Le modèle fonctionnel suivant résume et représente bien le principe de fonctionnement global d'un SAE basé sur le modèle OAIS :

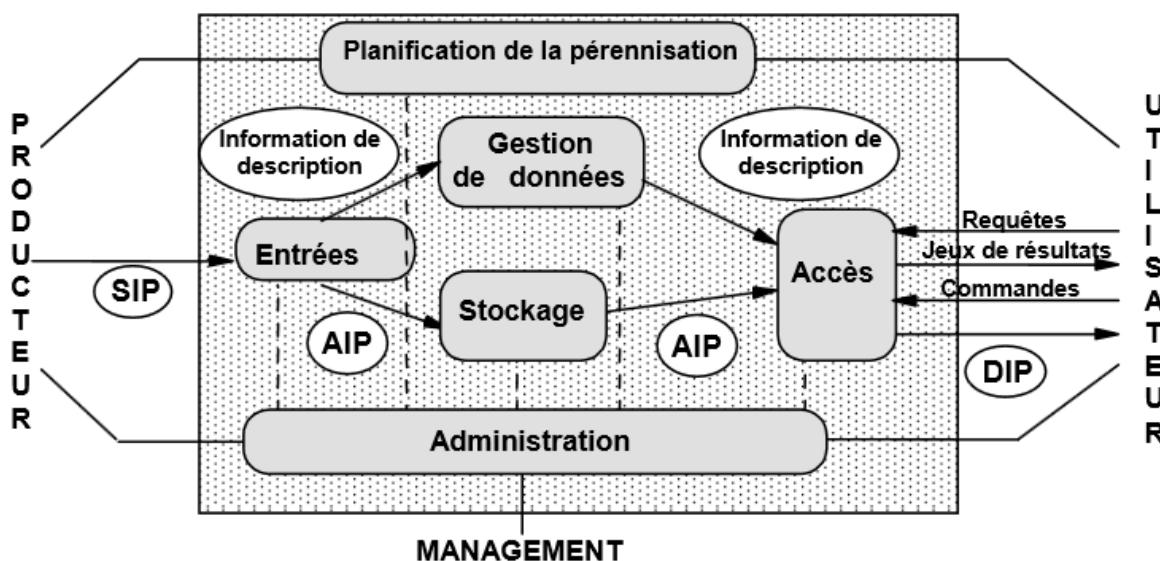


Figure 9 : Modèle fonctionnel OAIS¹

Un bon SAE devrait donc se conformer au modèle OAIS et respecter ce schéma.

¹ Schéma récupéré sur le document de la norme OAIS : http://pin.association-aristote.fr/lib/exe/fetch.php/public/documents/norme_oais_version_francaise.pdf

Précisons également qu'il existe une norme importante (**ISO 20652**) issue du standard **PAIMAS (Producer Archive Interface Methodology Abstract Standard)** qui s'applique à l'entité « Entrée » du modèle OAIS. Elle définit une méthodologie découpée en quatre phases pour les échanges entre le producteur et l'archive.

Son objectif est donc de spécifier les modalités de versement d'une archive conforme à la politique du SAE. Il fixe donc l'obligation au producteur de générer des SIP (données à verser) respectant les règles fixées afin que les AIP générées les respectent également.

PAIS (Producer-Archive Interface Specification) est quant à lui le standard d'implémentation de PAIMAS.

II.2.2.1 Normes et recommandations fonctionnelles pour l'archivage électronique

Plusieurs normes sont devenues de fait des exigences ou des recommandations pour la mise en œuvre d'un système d'archivage de documents électroniques. Ces normes d'archivage électronique découlent essentiellement de la norme **ISO 15489**.

Trois normes ou standards ressortent du paysage normatif pour les spécifications fonctionnelles et techniques d'un SAE. Il s'agit de la norme **NF Z042-013** (transposée en **ISO 14641-1**), et des standards **MoReq2010** et **ICA-Req** (devenu **ISO 16175**).

II.2.2.1.1 NF Z42-013 / ISO 14641-1

La norme **NF Z42-013** (Archivage électronique - Spécifications relatives à la conception et à l'exploitation de systèmes informatiques en vue d'assurer la conservation et l'intégrité des documents stockés dans ces systèmes), créée par l'AFNOR, définit les mesures techniques pour garantir l'intégrité, l'horodatage et la conservation des documents numériques.

Cette norme nationale est payante, elle a été traduite et transposée dans la norme internationale **ISO 14641-1**.

Elle assure entre autre la traçabilité des processus dans l'archivage électronique, telle que la numérisation ou l'horodatage. Initialement, elle limitait les supports de stockage aux seuls « WORM » (*Write Once Read Many*), c'est-à-dire aux supports logiques ou physiques non réinscriptibles. La nouvelle version de la norme autorise d'autres supports que le WORM, à savoir le support magnétique, à la condition d'avoir recours à la cryptographie ou la signature électronique afin de garantir l'intégrité du système [10]¹.

¹ Ballet P. & Bensoussan A. : *Une nouvelle norme pour l'archivage électronique*

La norme stipule que, pour prouver l'intégrité et la fidélité des documents électroniques et numérisés, l'organisme qui met en place un SAE doit réaliser périodiquement des audits du système ainsi que des procédures définies.

Pour les documents d'origine numérique, elle impose un ensemble de métadonnées à lier aux documents. Par exemple, l'heure d'enregistrement du document, le support physique utilisé, la marque et le numéro de série du matériel de stockage actuel, ainsi que le précédent...

Elle précise également entre autres les fréquences de sauvegarde, les opérations de migration à effectuer avant désuétude du matériel ou du format de données.

Un chapitre concerne la sécurité du système d'information du SAE, à savoir la sécurité d'accès aux locaux, les règles de sécurité définies pour le personnel... **[11]**¹

Enfin, la norme envisage le recours à des tiers-archiveurs en définissant alors les services et obligations attendus par le prestataire, tels que les normes, les méthodes de sécurité qu'il a mises en place ou encore la possibilité de réversibilité des archives.

II.2.2.1.2 MoReq2010

MoReq (*Modular Requirements for Records Systems ou Modèle d'Exigences pour l'Archivage Electronique*) a été développé par la Commission Européenne et le Document Lifecycle Management Forum Foundation (*DLM Forum Foundation*) en 2001. **MoReq2**² (février 2008), traduit en français par Marie-Anne Chabin³, ne traite pas de l'archivage patrimonial, mais est plutôt orienté archivage légal. MoReq2 a fortement été inspiré du modèle de métadonnées décrit dans la norme ISO 23081, norme fille de l'ISO 15489.

Il s'agit d'un recueil d'exigences pour la conservation des documents numériques, et plus précisément, il définit les fonctionnalités d'un SAE.

Il énonce des principes traitant du classement, de la sécurité, ou des métadonnées à définir pour l'archivage électronique.

Il est relativement détaillé pour chaque point à mettre en œuvre. Par exemple, pour la définition de processus de capture d'un document, il énonce 28 exigences et définit 39 exigences pour la mise en place d'un plan de classement.

Sa dernière version, **MoReq2010** qui a vu le jour en 2011, se veut plus abordable. Contrairement à l'ancienne version, **MoReq2010** se veut plus modulable, afin de séparer les

¹ Chabin M.A. : *Archivage Electronique : Problématiques et Normes*

² Version de MoReq2 traduite en français à l'adresse : <http://www.archivesdefrance.culture.gouv.fr/static/2094>

³ Entre autres professeur associé au CNAM et secrétaire générale du CR2PA (Club des responsables de politiques et projets d'archivage)

parties indispensables au fonctionnement d'un SAE et celles qui peuvent être mise en œuvre selon les spécificités et les besoins du système.

Il définit également une architecture du SAE découpée en services, comme représentée sur la figure ci-dessous.

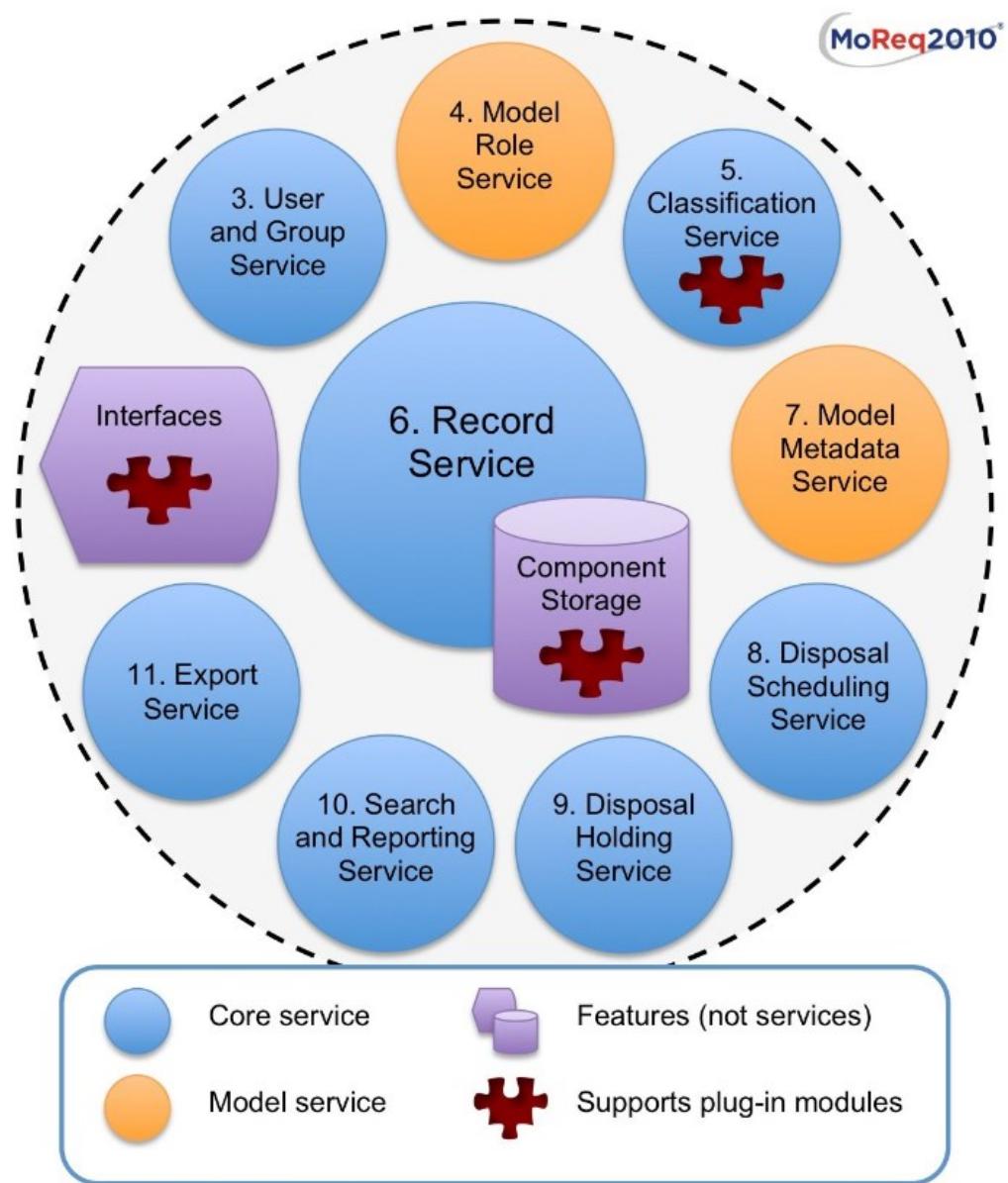


Figure 10 : Architecture découpée en services d'un SAE conforme MoReq2010¹

¹ DLM Forum Foundation : image récupérée sur les spécifications de MoReq2010, accessible à l'adresse : <http://moreq2010.eu/pdf/MoReq2010-Core+Plugin%28v1-0%29.pdf>

II.2.2.1.3 ICA-Req / ISO 16175

Il s'agit d'un standard développé par le Conseil International des Archives (CIA). **ICA-Req**¹ définit les principes et les exigences fonctionnelles pour la mise en place d'un archivage électronique. **ICA-Req** a été normalisé et est devenu la norme **ISO 16175**.

Cette norme est composée de trois modules. Le premier présente le contexte et les principes, le deuxième aborde les recommandations et les exigences pour la mise en place d'un SAE et est proche de MoReq en étant moins détaillé.

Enfin le dernier module traite des notions de documents électroniques dans des applications métiers, approches utiles pour des organismes ne disposant pas de SAE.

Elle s'adresse par conséquent en partie aux organismes souhaitant mettre en place un SAE mais également aux organismes mettant en place des systèmes d'informations dont le contenu est susceptible d'intégrer à terme un SAE **[12]**².

Les différentes normes et standards présentés jusqu'ici permettent de mettre en place une politique d'archivage avec les bons processus, ainsi que de modéliser les flux du SAE et de définir les spécificités fonctionnelles et techniques auxquelles il devra répondre.

II.2.2.2 Normes couvrant les métadonnées

Comme nous l'avons déjà mentionné, les métadonnées permettent de décrire, de contextualiser, de classer un document ou encore de gérer sa traçabilité.

Elles sont essentielles pour la bonne compréhension future de l'archive et sa lisibilité. Elles jouent donc un rôle crucial dans le processus de pérennisation. Elles sont en outre un outil d'aide à la recherche dans un système d'archivage.

Les métadonnées peuvent être caractérisées de différentes manières. Par exemple, les métadonnées contenant les dates de conservation sont plus liées à la pérennité, alors que celles traitant de l'encodage ou du format d'un document sont plutôt techniques. Elles peuvent également être administratives, contextuelles, structurelles...

¹ Version d'ICA-Req traduite en français à l'adresse suivante : www.ica.org/download.php?id=1115

² Direction des Archives de France : ICA-Req

Par conséquent, elles sont aussi importantes que les données de contenu en elles-mêmes et doivent donc être conservées au même titre dans un système d'archivage.

Il convient alors de prévoir un langage, un socle commun entre les archives et les producteurs d'archives. Beaucoup de normes existent en la matière pour répondre à ce besoin.

II.2.2.2.1 Normes archivistiques

Parmi les normes plutôt axées sur le métier de l'archiviste, nous pouvons citer l'**EAD DTD** (*Encoded Archival Description ou Description Archivistique Encodée*), l'**EAC-CPF** (*Encoded Archival Context - Corporate Bodies, Persons and Families ou Contexte Archivistique Encodé - Collectivités, Personnes et Familles*) ou encore **Dublin Core**. Elles sont toutes les trois basées sur des XSD (schéma XML).

L'**EAD DTD** permet de structurer les descriptions des documents d'archives, et donc de standardiser les instruments de recherche des archives. Il fonctionne en corrélation avec la norme l'**ISAD(G)**¹ en permettant de structurer les descriptions définies par cette dernière. C'est devenu un format courant de catalogage en France, principalement dans les services d'archives, les musées ou les bibliothèques [15]².

La norme **EAC-CPF** permet la description des producteurs d'archives. Elle est conforme à la norme **ISAAR-CPF**³.

Enfin le standard **Dublin Core** (transposé dans la norme **ISO 15836**) est une norme beaucoup utilisée par les bibliothèques, mais également les archives. Il permet une description thématique et également dispose de balises pour la gestion du droit d'auteur. On la retrouve souvent dans les outils de GED et de SAE.

II.2.2.2.2 Normes généralistes

Ces différentes normes permettent de gérer des métadonnées de domaines différents, afin d'augmenter la lisibilité future des archives. D'autres normes plus généralistes de description

¹ Norme générale et internationale de description archivistique (International Standard Archival Description-General)

² Bibliothèque Nationale de France : EAD - Encoded Archival Description

³ International Standard Archival Authority Record for Corporate Bodies, Persons and Families (Norme Internationale sur les Notices d'Autorité utilisées pour les Archives relatives aux Collectivités, aux Personnes et aux Familles)

des archives existent, il s'agit de **METS** et **PREMIS**, toutes deux maintenues par la Bibliothèque des Congrès¹.

II.2.2.2.2.1 PREMIS

PREMIS (*PREServation Metadata: Implementation Strategies*) est un standard permettant de gérer une partie des données de préservation (ou de pérennisation au sens OAIS). Par exemple, pour qu'un document reste viable à longue échéance, il devra contenir des informations décrivant les dates de modifications des autres métadonnées, ou encore bien décrire la ou les signatures électroniques de l'archive.

Il ne gère pas toutes les métadonnées descriptives de préservation, comme par exemple celles qui seraient liées uniquement à un format spécifique, ou encore à un métier. Mais il représente un sous-ensemble des métadonnées de pérennisation d'une archive.

Un exemple d'implémentation est donné en annexe (Voir Annexe 1 : Exemple PREMIS).

II.2.2.2.2.2 METS

METS (*Metadata Encoding and Transmission Standard*) est également un standard basé sur un XSD maintenu par la bibliothèque du congrès aux USA (*Library of Congress*), tout comme **PREMIS**. Le document XML (*eXtensible Markup Language*) généré permet de structurer de manière claire les différentes informations décrivant un document d'archive.

Au sein du même fichier, il peut par exemple intégrer dans des zones définies des métadonnées de différentes natures, telles que **EAD**, **Dublin Core** ou encore **PREMIS [17]**². Étant donné qu'il permet d'associer ces métadonnées, ce conteneur représente un véritable avantage. Il est possible de référencer également les métadonnées vers des liens extérieurs, afin de ne pas dupliquer les métadonnées techniques identiques à plusieurs fichiers.

Il définit « un cadre pour le stockage de métadonnées administratives, structurales et descriptives pour les documents numériques [18]³ ». Il permet de grouper les fichiers constituant l'archive d'être regroupés dans des sections grâce à l'utilisation de balises particulières.

Il est complémentaire du modèle conceptuel OAIS du fait qu'il permet de structurer les paquets d'information AIP, SIP et DIP. **METS** doit être vu comme un guide à la mise en place

¹ Library of Congress : <http://www.loc.gov/standards>

² Direction des Archives de France : Métadonnées

³ Library of Congress : METS - Metadata Encoding & Transmission Standard

d'un SAE, dans l'optique de permettre la gestion des paquets d'informations au sens OAIS et particulièrement leur versement dans un autre SAE (par exemple les AIP lors d'une migration).

Un exemple de document **METS** est disponible en *Annexe 2 : Exemple METS*

II.2.2.3 Normes pour l'échange de données

II.2.2.3.1 Le SEDA

Le **RGI** préconise un standard pour l'échange de données entre administrations, mais également pour les échanges entre les différents acteurs de l'archivage. Il s'agit du **SEDA** (*Standard d'Échange de Données pour l'Archivage*).

« *Le standard d'échange de données pour l'archivage modélise les différentes transactions qui peuvent avoir lieu entre des acteurs dans le cadre de l'archivage de données. Ces transactions sont au nombre de six : le transfert, la demande de transfert, la modification, l'élimination, la communication et la restitution. Les acteurs sont eux au nombre de cinq : le service producteur, le service versant, le service d'archives, le service de contrôle et le demandeur d'Archives* » [19]¹.

Le **SEDA** définit les scénarios à mettre en œuvre pour chaque transaction, sous forme de **diagrammes UML**, la forme et le contenu des messages sous forme XML que peuvent ou doivent s'échanger les acteurs.

Il est **décliné en profils**, selon le type de données que l'on souhaite transmettre à un service d'archive, ou selon l'application métier qui doit transmettre un document au SAE. L'Adullact propose d'ailleurs un outil graphique pour générer des profils².

Afin de ne pas être supplanté par une future norme internationale, le **SEDA** a été construit de manière à rester conforme aux standards ou normes tels qu'UML, XML, OAIS, EAD DTD ou encore PREMIS.

Depuis janvier 2014, le **SEDA** a été normalisé par l'AFNOR sous la norme **Z44-022** ou **MEDONA** (*Modélisation des Échanges de DONnées pour l'Archivage*).

¹ Direction des Archives de France (DAF) : *Le standard d'échange de données pour l'archivage (SEDA)*

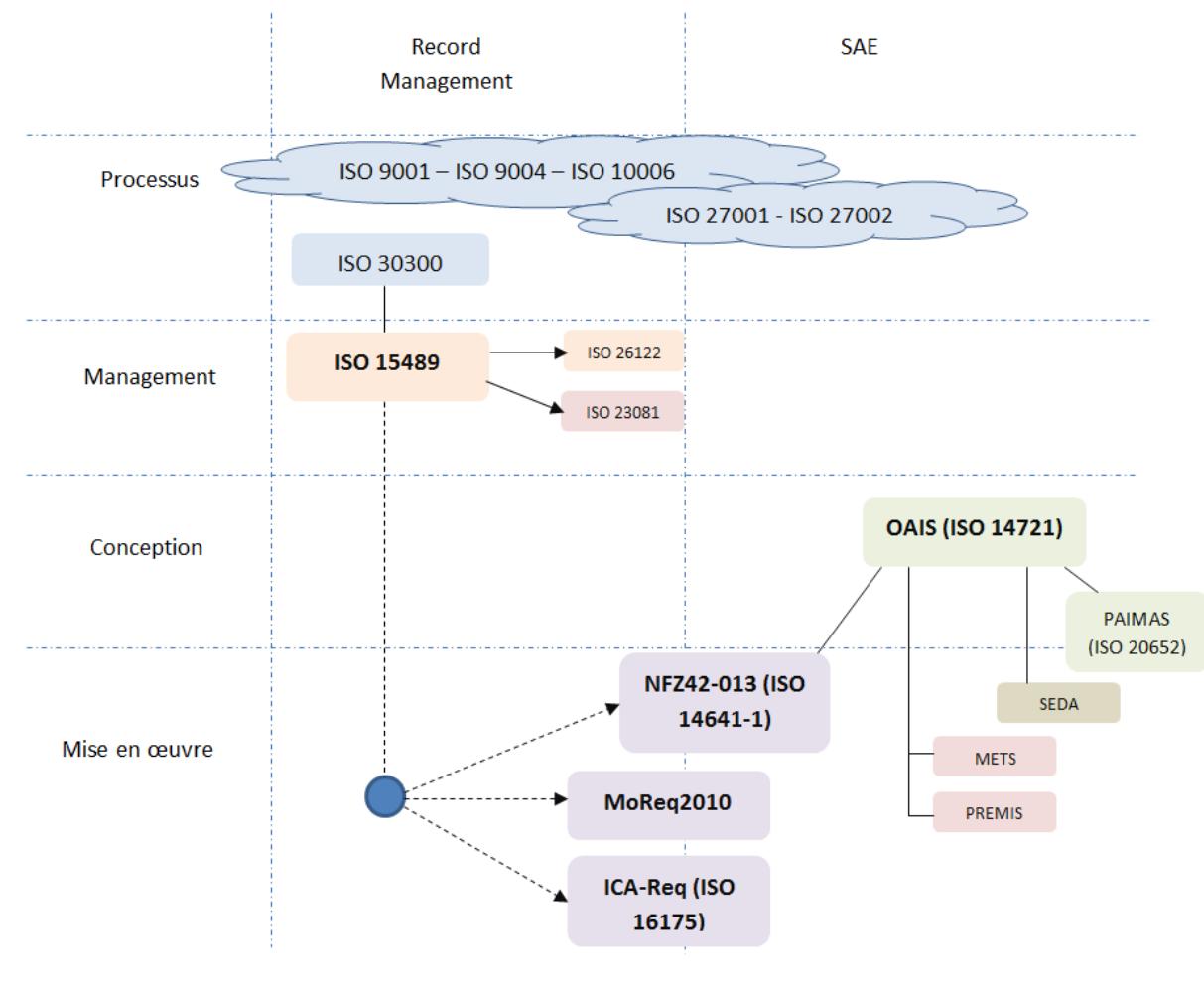
² AGAPE (de Application de Génération Automatisée Profils Électroniques) de l'Adullact : <http://agape.adullact.net/>

II.2.2.4 Tableau récapitulatif

Il est facile de se perdre dans ce vaste paysage normatif, les normes couvrant l'archivage électronique sont légions. Pour s'y retrouver, il faut découper l'intégration de ces normes à différentes étapes d'un projet de mise en place d'un SAE.

Le tableau suivant récapitule l'ensemble des normes abordées (tableau inspiré des schémas de Marie-Anne Chabin¹ et Arnaud Hulstaert²).

Tableau II : Tableau récapitulatif des normes en matière d'archivage électronique



¹ Schéma de Marie-Anne Chabin accessible à l'adresse : <http://www.archive17.fr/index.php/Normes-et-guides/normes-et-guides-records-management-preservation.html>

² Schéma d'Arnaud Hulstaert accessible à l'URL : <http://blogresearch.smalsrech.be/?p=3284>

II.3 Formats et supports de stockage préconisés pour la pérennisation

Outre les normes et standards permettant de faciliter ou de diriger la mise en place d'un SAE, il est bon de se référer aux bonnes pratiques concernant les supports de stockage numériques, ainsi que ceux traitant des formats de documents. En effet il faut prendre conscience de la fragilité de la conservation des documents numériques.

Un document numérique à conserver est stocké sur un support particulier dans un format particulier. Il a été généré par un logiciel, dans une version particulière, dans un encodage et un format particulier. Il peut également être conservé sur différents types de support, tels qu'un CD, un DVD, un disque dur, une bande magnétique, le Cloud...

Comme nous l'avons vu précédemment, les métadonnées de pérennisation permettent de stocker ces informations, à savoir comment le document a été créé et quels ont été ses différents supports de stockage.

Toutefois, afin de minimiser les risques de perdre les données contenues dans le document numérique, ainsi que les risques de le voir devenir illisible, il est utile de respecter certaines bonnes pratiques en vigueur.

II.3.1 Encodage

Comme nous l'avons déjà dit, un fichier de type texte par exemple, peut être encodé dans différents jeux de caractères. Cette différence provient de la multitude des langues, et surtout des alphabets existants. Par exemple, les accents de la langue française n'existent pas dans la langue anglaise. Un fichier de langue française encodé avec le jeu de caractère américain ASCII (*American Standard Code for Information Interchange*) serait donc incomplet, du fait de l'absence des accents.

Il y a peu de chance que les jeux de caractère existants à l'heure actuelle disparaissent, mais autant que possible, il est préférable d'utiliser un format d'encodage qui couvre l'ensemble des langues.

C'est le cas du standard UNICODE (ISO 10646), ainsi que de l'UTF-8 (*Universal character set Transformation Format – 8 bits*) qui en découle et permet de coder l'ensemble des caractères universels.

Il est donc préconisé d'utiliser cet encodage pour les données contenant des chaînes de caractères. Il est également conseillé d'utiliser en ce qui concerne les éditeurs de texte, des polices de caractères libres telles qu'Arial ou Times New Roman.

II.3.2 PrésERVER les documents en fonction des formats utilisés

Afin de pouvoir être lu, il faut avoir connaissance du format d'un fichier numérique. En effet, le logiciel nécessaire à l'ouverture d'un fichier de type audio, vidéo, image ou encore texte sera différent selon chaque cas. De plus, même des fichiers de même type peuvent être de différents formats.

Un fichier image pourra par avoir comme extension *.jpeg*, *.tiff*, *.png* ou encore *.gif*. Un logiciel permettant d'ouvrir les fichiers images ne pourra pas forcément gérer tous les types de fichiers images.

De plus, le fichier peut avoir été généré dans une version particulière d'un logiciel, et par conséquent, même avec une version plus récente du même logiciel, il n'est pas garanti que l'on puisse ouvrir le fichier ou du moins préserver son intégrité. Un exemple courant est celui des fichiers du traitement de texte de Microsoft Word. Un fichier Word (.doc) généré avec Word 97 risque en effet de voir au minimum sa mise en forme modifiée si on l'ouvre avec une version de Word plus récente, Word 2010 ou 2013 par exemple. Enfin, il n'est pas garanti que le logiciel en question, même dans la version qui a servi à générer le document, fonctionne correctement dans quelques années par exemple sur les futurs systèmes d'exploitation.

À titre d'exemple, voici un extrait de texte écrit et visualisé avec Word 2010 :

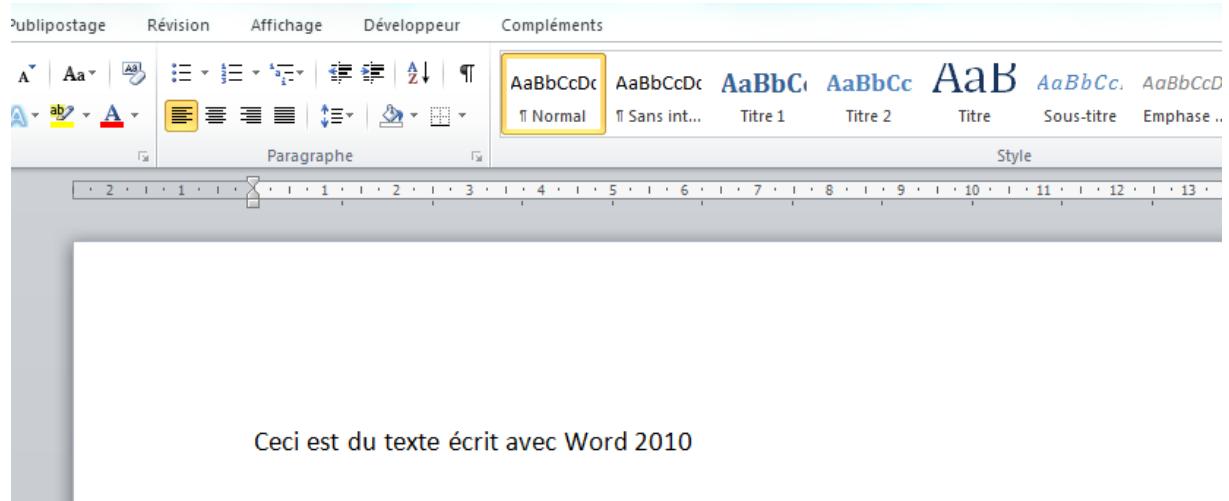


Figure 11 : Exemple de texte visualisé dans Word 2010

Et voici le même document ouvert avec l'éditeur de texte Notepad++ :

Figure 12 : Exemple d'un document Word 2010 visualisé avec Notepad++

Il est illisible avec ce dernier logiciel.

II.3.2.1 Formats préconisés

Dans le cadre de l'archivage électronique, il faut avoir mis en place des règles fixant les formats autorisés au sein du système, comme l'a fait le CINES (*Centre Informatique National de l'Enseignement Supérieur*) sur sa plateforme d'archivage PAC (*Plateforme d'Archivage du CINES*) [20]¹.

Ces formats doivent être choisis parmi des formats normalisés ou du moins standardisés, et de préférence des formats ouverts et non propriétaires. Par conséquent, l'exemple précédent du document Word ne devrait pas pouvoir se produire, car il faudrait prohiber ce genre de format dans un tel système.

Le RGI liste tout un ensemble de formats préconisés pour l'échange de données entre les collectivités. Pour la mise en place d'un système d'archivage électronique, il peut être bon, voire obligatoire, de s'y référer et de choisir les formats qui seront autorisés dans le SAE parmi ceux-ci.

¹ CINES : Liste des formats archivables par la plateforme PAC

Nous présentons ici quelques formats pouvant être considérés comme pérennes à travers différents types de données, tels que les fichiers texte, audio, vidéo ou encore de bases de données.

II.3.2.1.1 Texte

Pour les fichiers de type texte, il faut privilégier le format PDF¹/A1 (norme ISO 19005-1) qui gère très bien les documents avec mise en page, images ou encore avec des polices de caractères spécifiques. Un fichier au format brut.txt pourrait également être accepté dans le système.

II.3.2.1.2 Audio

Concernant les fichiers audio, s'il s'agit d'un enregistrement haute qualité, il est préconisé d'utiliser le format WAV sans compression, alors que du MP3 (basé sur le standard MPEG-1 devenu la norme ISO 11172-3), suffira pour de l'audio basse qualité.

II.3.2.1.3 Vidéo

Pour la vidéo, comme pour l'audio, le RGI préconise de conserver dans deux formats distincts en fonction de la qualité de la vidéo. Pour une vidéo HD (*Haute Définition*), le standard MPEG-4 est recommandé, alors que le standard MPEG-2 l'est pour la vidéo basse définition.

II.3.2.1.4 Images

Pour les fichiers de type images, il est recommandé des standards tels que JPEG, TIFF, PNG ou encore GIF.

II.3.2.1.5 Bases de données

Le RGI préconise l'utilisation du XML ou du CSV (*Comma-Separated Values*) pour l'export des bases de données. Toutefois, la DISIC (*Direction Interministérielle des Systèmes d'Information et de Communication*) conseille le SIARD (*Software Independent Archiving of Relational Databases*) [20]², format de bases de données relationnelles développé par les AFS (*Archives Fédérales Suisses*).

Cet outil, disponible gratuitement, permet d'archiver plusieurs types de SGBDR (*Système de Gestion de Bases de Données Relationnelles*) telles qu'Oracle, MySQL ou encore SQL Server.

¹ PDF : Portable Document Format

² DISIC : Archivage électronique - Un nouveau domaine d'expertise au service de la gouvernance des systèmes d'information

II.3.2.2 Validation des formats

Le fait qu'un **fichier soit dans un format particulier** n'implique pas forcément qu'il soit **conforme aux spécifications dudit format**. Afin d'évaluer sa conformité au format, il existe des outils de validation.

Le **CINES** met à disposition **un outil disponible gratuitement** qui valide un panel de formats. Il s'agit de l'outil **FACILE (outil de validation du Format d'Archivage du CINES par anaLyse et Expertise)**, accessible à l'adresse : <http://facile.cines.fr/>. Les étudiants doctorant doivent d'ailleurs valider les fichiers composant leur thèse dès leur création, afin qu'ils soient bien conformes au système d'archive PAC.

II.3.3 Supports de stockage

Le problème est un peu similaire concernant les supports de stockage. Le support quel qu'il soit peut en effet s'être dégradé au fil des ans, empêchant alors la lecture de son contenu. Ceci peut être du fait de différents facteurs tels que le taux d'humidité, la qualité du support, la présence de poussières...

Autre problème possible, le support utilisé pour la conservation peut avoir été très bien conservé, mais faut-il encore qu'il existe les lecteurs permettant de les lire. Un exemple parlant à tout le monde, sauf aux plus jeunes, est celui des cassettes VHS. Même si la VHS est bien conservée, il est devenu très difficile de trouver des magnétoscopes pour les lire.

II.3.3.1 Types de supports de stockage

Les supports de stockage peuvent être de différents types. Les plus couramment utilisés sont les supports optiques (CD-R, DVD ou encore Blu-Ray) et les supports magnétiques (disques durs et bandes). Des supports de type magnéto-optiques existent également, mais ils coûtent relativement cher.

Dans le cadre de l'archivage et afin de préserver l'intégrité de documents, il est bon de choisir **des supports WORM physiques ou logiques**. Un CD réinscriptible (CD-RW) ne sera en effet pas possible. Un support de type WORM logique mettra en œuvre de manière logicielle le blocage **empêchant la modification des documents**.

Le choix d'un support est **lié à une analyse de la volumétrie des données à archiver**. On ne va en effet pas stocker des téraoctets (*To*) de données sur des CD, DVD et même Blu-Ray. On optera alors pour les supports tels que les **baies de disques et les bandes magnétiques**.

Pour les gros volumes, le **Cloud (informatique en nuage)** pourrait apparaître la solution idéale. Dans ce cas, ce n'est plus à l'organisation de gérer directement son espace de stockage, mais directement au fournisseur d'une telle solution. Toutefois les législations étant différentes d'un pays à l'autre, il est déconseillé de stocker ses données d'archivage dans des pays hors de l'Union Européenne. De plus, s'il s'agit de données médicales, les hébergeurs doivent obtenir un agrément auprès du ministère de la santé, qui ne sera pas délivré pour des datacenters (entrepôt de données) situés hors du territoire français. Concernant l'archivage définitif, la question ne se pose pas à l'heure actuelle, étant donné que les données doivent rester géographiquement dans les locaux de la collectivité.

Outre le Cloud, des solutions de type **NAS (Network Attached Storage)** ou **SAN (Storage Area Network)** avec **des technologies RAID (Redundant Arrays of Inexpensive Disks)** peuvent être envisageables. Elles permettront de sécuriser (un peu) les pannes matérielles.

Concernant le stockage sur les supports optiques (CD, DVD et Blu-Ray), diverses études de durée de vie selon les marques, les graveurs utilisés ou encore la vitesse de gravure ont été effectuées par le LNE (*Laboratoire National de métrologie et d'Essais*). Elles sont accessibles sur le site des Archives de France, à l'adresse suivante :

<http://www.archivesdefrance.culture.gouv.fr/gerer/archives-electroniques/formats-et-supports/>

Avec ce type de support, il est conseillé de doubler les gravures, mais également de re-graver les archives bien avant la date prévue, avant que le support ne se détériore.

Différentes études menées par le LNE ont fait ressortir le tableau suivant, montrant les estimations de durées de vie des différents types de supports de stockage :

Tableau III : Estimation de la durée de vie des supports d'archivage électronique

-	Estimations des fabricants	Estimations issues d'études scientifiques
CD-R	Entre 100 et 300 ans	Entre 1 et 15ans
DVD-R	50 ans	Entre 1 et 5 ans
Bandes magnétiques	Entre 10 et 30 ans	Entre 1 et 5 ans
Disque dur	Entre 5 et 10 ans	Entre 3 mois et 7 ans

Comme on peut le voir sur ce tableau, les supports de stockage actuels n'ont qu'une durée de vie très limitée, bien en dessous de celles annoncées par les fabricants.

II.3.3.2 **Supports prometteurs**

Outre ces supports traditionnels, les chercheurs travaillent à de nouveaux support de stockage, plus pérennes, plus résistants ou permettant encore de stocker beaucoup plus de données. Ces études deviendront peut-être des pistes futures pour les choix à effectuer en matière d'archivage de l'information numérique.

Voici quelques-uns de ces supports que nous trouvons intéressant de présenter ici.

Une technologie intéressante est celle du **disque en saphir** synthétique, nommée la **nanoforme**. Une jeune société française, Arnano a créé ce système qui offre un support très résistant ne nécessitant pas ou peu de précautions pour la conservation à très long terme. Les informations n'étant pas codées, il n'y aurait donc pas à se soucier des évolutions technologiques. En effet, la technologie se rapproche des technologies microfilms, à savoir, les documents sont stockés comme des images miniatures sur le disque, le disque pouvant stocker 10000 pages A4.

La couche en saphir les rend inaltérables dans le temps, résistante à des températures très élevées (1700 °C). Selon Alain Rey, le fondateur et inventeur de la technologie, les disques pourraient rester pérennes sur plusieurs millénaires. Il s'agit véritablement de support WORM, donc non modifiable, garantissant par conséquent l'intégrité des documents et permettant de conserver les valeurs probantes.

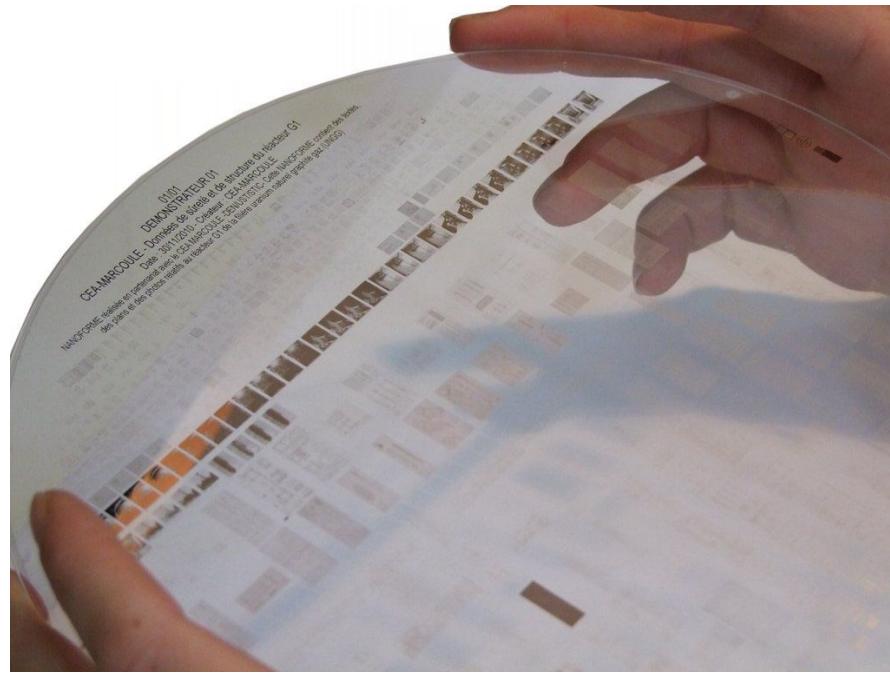


Figure 13: Disque nanoforme de la société Arnano¹

¹ Image récupérée à l'URL suivante : <http://www.futura-sciences.com/magazines/high-tech/infos/actu/d/informatique-disque-arnano-archive-donnees-siecles-39638/>

D'un autre côté, des chercheurs anglais de l'université de Southampton ont réussi à stocker dans un **disque de verre** de quelques millimètres, via des mécanismes nanométriques, 50 Go de données. Outre la capacité de stockage, le verre offre également des qualités de résistance élevée, un peu moindre que le saphir, mais permettant par exemple de pouvoir résister à des températures élevées (plus de 900 °C) et d'être conservé sur de très nombreuses années. Contrairement à la technologie nanoforme, ces supports sont réinscriptibles [22]¹.

Sur un principe similaire, mais avec du diamant et non du verre, ont été créés des disques pouvant atteindre les mêmes capacités qu'un Blu-Ray, à savoir 25 Go.

D'autres supports minéralogiques sont également proposés à la commercialisation ou le seront dans les prochaines années, comme le M-Disc (technologie sur diamant développée par Millenniata et LG) ou encore des solutions de stockage en quartz [23]².

Enfin, dans un domaine un peu différent, de chercheurs ont réussi à stocker des données numériques dans de l'ADN ou encore dans des chaînes moléculaires synthétiques [24]³. Outre la conservation à long terme, le principal avantage serait les capacités exceptionnelles de stockage. Plus de 5 Mo de données avaient déjà été stocké dans un milliardième de gramme d'ADN. Selon Sriram Kosuri, un des chercheurs, l'ensemble des données numériques mondiales (1,8 Zo, un zettaoctet valant 10^{21} octets) pourrait tenir sur 4 grammes d'ADN [25]⁴ !

II.3.3.3 Conseils généraux pour la pérennité des supports de stockage

En attendant ces futures technologies qui permettront probablement d'offrir une pérennité plus longue et plus solide, il est possible de respecter quelques bonnes pratiques afin de préserver au mieux les documents numériques.

Tout d'abord, quel que soit le support choisi pour l'archive, les premières précautions à respecter concernent l'environnement du stockage. Il ne doit pas être humide, trop chaud ou trop froid, poussiéreux ou encore soumis à des perturbations magnétiques.

Il convient également de **doublonner les supports par des mécanismes de réPLICATIONS, de sauvegardes, qu'il faudra conserver sur des sites géographiques distants.**

¹ Site de gurumed : Les cristaux de mémoire pourraient devenir une réalité

² Bastide, A. : Le quartz, support de stockage pour des milliers d'années

³ Cailloux, L : Des molécules pour stocker l'information

⁴ Cheminat, J : L'ADN, une solution pour le stockage de demain

Enfin, il est conseillé de varier les types de support pour la conservation numérique, ou au moins varier les marques du support comme le préconise le GIS-SPADON [22]¹. C'est-à-dire qu'il faut par exemple stocker les données sur des baies SAN (disques durs), et en parallèle les stocker également sur des bandes magnétiques.

II.3.4 Migrations

Afin de conserver la lisibilité et donc la pérennité des documents archivés, il faut prévoir des migrations de support et de format.

Comme nous l'avons déjà dit, les supports de stockage actuels ont une durée de vie très limitée, surtout lorsqu'on parle d'archivage patrimonial, et donc de conservation sur plusieurs centaines d'années. Il faut alors prévoir le transfert des données de préférence avant les durées estimées de fin de vie des supports (*cf. Estimation de la durée de vie des supports d'archivage électronique*) vers des supports plus récents.

Si une technologie utilisée pour le stockage est en voie de disparition prochaine, il faut impérativement effectuer des migrations vers d'autres types de support, en veillant bien à ne pas modifier le contenu des données.

Concernant les formats de fichiers, le problème est le même. Si certains formats utilisés par le système d'archivage sont de moins en moins utilisés et sont donc amenés à disparaître, il faut également prévoir des migrations du format d'origine vers des formats plus pérennes.

Ce tour d'horizon offre la possibilité d'effectuer des choix judicieux lors de la mise en place d'un système d'archivage électronique. Les normes présentées apporteront un cadrage méthodologique, les préconisations en matière de formats de fichiers et de supports de stockage permettront de réduire les difficultés rencontrées dans un tel projet.

Cet état de l'art va nous permettre de spécifier dans la prochaine partie un démonstrateur et de définir les fonctionnalités qu'on doit en attendre.

¹ *GIS-SPADON (ex GIS-DON) Groupement d'Intérêt Scientifique sur les Supports Pérennes d'Archivage des DOnnées Numériques : Recommandations du GIS-DON concernant la préservation à long terme des données numériques*

III Spécifications du démonstrateur

Avant la réalisation et la mise en œuvre d'un prototype de briques d'archivage patrimonial, il convient de spécifier les fonctionnalités fonctionnelles et techniques attendues par le démonstrateur.

C'est ce que nous allons faire dans cette partie, en présentant tout d'abord le concept de SAE à travers ses fonctionnalités. Après avoir identifié les fonctionnalités de la maquette, nous effectuerons une petite étude de marché afin de prendre connaissance des solutions existantes sur le marché et des critères qui nous permettront de définir la solution retenue.

Dans son fond documentaire, l'ANSSI fournit un exemple de cahier des charges [26]¹ afin d'accompagner les collectivités territoriales dans le choix d'une solution d'archivage électronique. Ce document nous a servi de base pour la définition des spécifications.

III.1 SAE

Un Système d'Archivage Électronique est un ensemble d'infrastructures matérielles et logicielles permettant la **conservation et la restitution de documents électroniques archivés**.

Il doit offrir les mêmes garanties que celles décrites pour le record management (**traçabilité, intégrité, sécurité et pérennité**). Il doit également faciliter la recherche de documents d'archives, leur exploitation et leur communication.

Enfin, un SAE doit impérativement permettre la réversibilité de son contenu. En effet, il est probable et même certain que le système ait une durée de vie plus courte que des archives patrimoniales.

III.1.1 Distinctions SAE, GED et CFN

Il existe également des confusions terminologiques entre la définition d'un SAE, d'une GED ou encore celle d'un coffre-fort numérique (CFN).

Une GED traite entre autres des données courantes, permet de les classer, de les rechercher. L'auteur d'un document peut le supprimer, il peut également exister plusieurs versions du même document, à l'inverse d'un SAE.

¹ Agence Nationale de Sécurité des Systèmes d'Information : Cahier des charges pour un système d'archivage électronique

Un CFN n'intervient que sur le périmètre de la conservation et l'intégrité des données archivées. Il ne tient pas compte de spécificités liées au Code du Patrimoine ou aux processus métiers liés à l'archivage comme le ferait un SAE. En particulier, contrairement à un SAE, un CFN ne dispose pas de workflows de versement, de workflows de communicabilité, de workflows d'élimination...

Voici ci-dessous un tableau présentant un comparatif de certaines fonctionnalités offertes par un SAE, une GED et un CFN.

Tableau IV : Fonctionnalités comparées SAE / GED / CFN

Fonctionnalités	SAE	GED	CFN
Modifications du document	Non	Oui	Oui
Coexistence de différentes versions	Non	Oui	Oui
Suppression par le créateur	Non	Oui	Oui
Traçabilité	Oui	Non	Oui
Procédure spécifique pour suppression (workflow)	Oui	Non	Non
Contrôle communicabilité	Oui	Non	Oui
Authentification forte	Oui	Oui	Oui
Contrôle intégrité	Oui	Non	Oui
Gestion des volumes de stockage	Oui	Oui	Non
Migration des formats	Oui	Non	Parfois
Migration des supports de stockage	Oui	Oui	Non
Modification des métadonnées	Oui	Oui	Parfois

III.2 Spécifications fonctionnelles et techniques

Nous proposons ici de présenter les fonctionnalités fonctionnelles et techniques auxquelles devra répondre la maquette du SAE.

Le SAE devra se baser sur le modèle OAIS présenté dans la partie précédente de ce document. Il devra donc comporter les entités/fonctions « Versement (ou Entrées) », « Gestion des données », « Planification de la pérennisation », « Stockage », « Accès / Consultation » et « Administration ».

Pour rappel, nous remettons le schéma présentant le modèle fonctionnel OAIS comprenant les différentes fonctions que devra offrir la maquette.

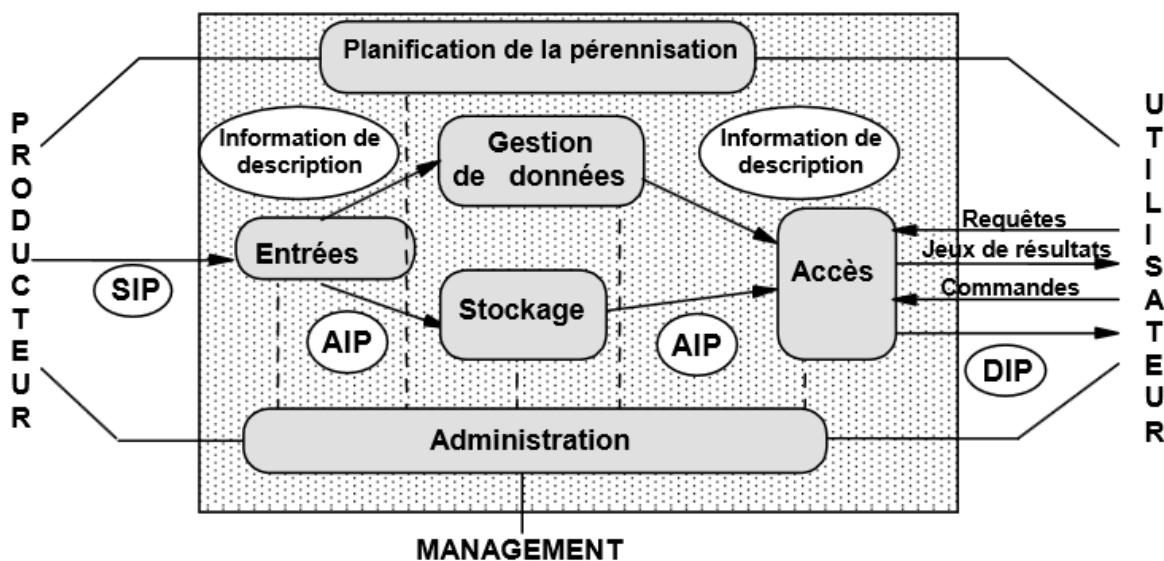


Figure 14 : Modèle fonctionnel OAIS¹

¹ Schéma récupéré sur le document de la norme OAIS : http://pin.association-aristote.fr/lib/exe/fetch.php/public/documents/norme_oais_version_francaise.pdf

Avant de présenter les fonctionnalités, nous pouvons d'ores et déjà positionner le futur système dans son environnement. Ceci peut être représenté par le diagramme de contexte suivant :

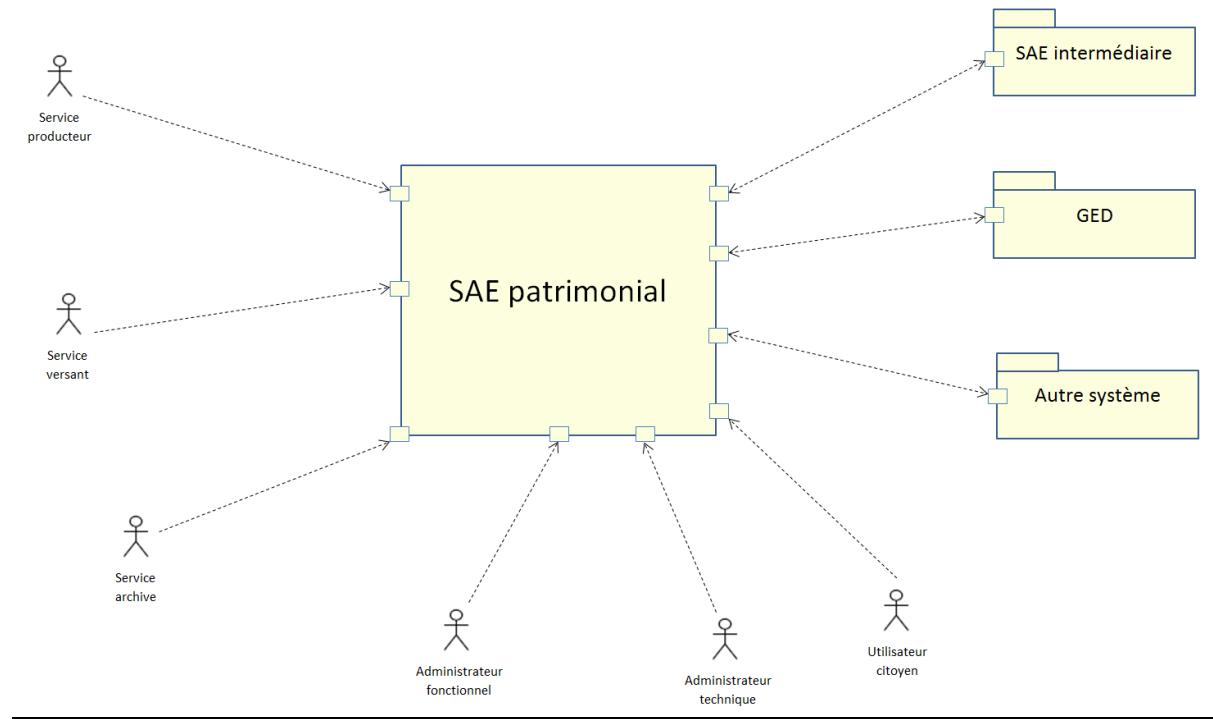


Figure 15 : Diagramme opérationnel de contexte du système de SAE patrimonial

Nous pouvons également positionner les différents acteurs du SAE autour des principales fonctionnalités d'un SAE. Le diagramme des tâches suivant peut représenter cette information :

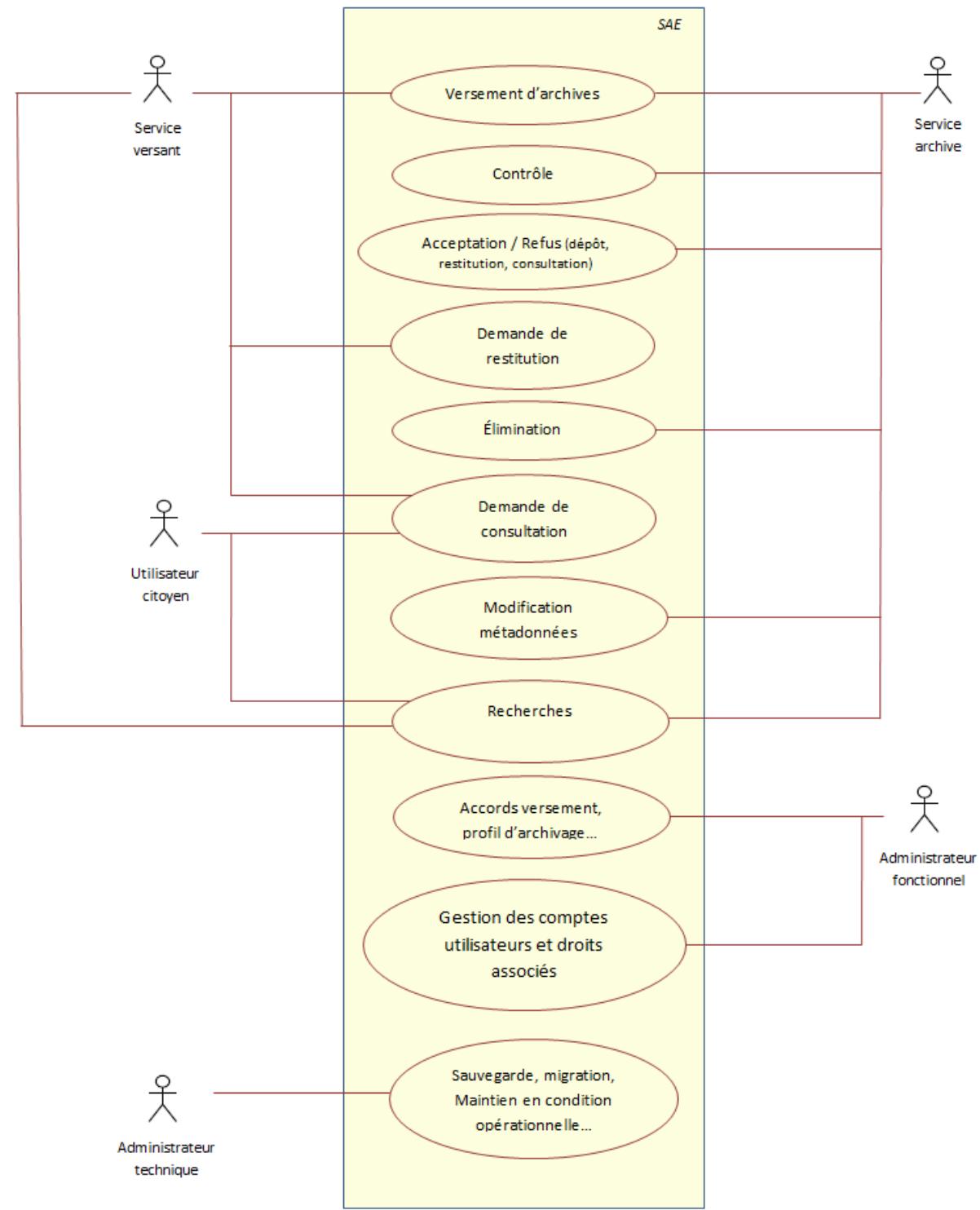


Figure 16 : Diagramme des tâches des différents acteurs du SAE

III.2.1 Spécifications fonctionnelles

Comme nous l'avons dit précédemment, et afin de définir les spécifications fonctionnelles, nous allons nous baser sur le modèle OAIS. Plus précisément, nous pouvons axer le découpage des fonctionnalités sur les entités fonctionnelles du modèle.

III.2.1.1 *Entité Versement (ou Entrée)*

La première finalité du SAE est de conserver les documents archivés. Pour cela, il convient de lui verser les futures archives.

Par conséquent, le SAE devra pouvoir recevoir des paquets d'informations à verser (*SIP*). Ce transfert devra pouvoir se faire soit directement sur la future plateforme, ou alors via l'appel d'API (*Application Protocol Interface*).

Le système devra être capable de vérifier l'intégrité des informations reçues (aucune perte de données). Les paquets transférés devront respecter le standard *SEDA*.

De plus, le transfert devra être conforme au « contrat » (accord de versement) défini entre les services versants et le service d'archives du système. Celui-ci définira entre autres les données descriptives obligatoires, les formats de fichiers autorisés, la structuration globale des paquets. Il sera donc possible de valider ou rejeter une demande de versement.

Le système devrait, si possible, être capable de convertir des documents qui ne seraient pas dans un format conforme. Cette transformation de format ne concerne bien entendu pas l'ensemble des formats existants mais plutôt des formats de fichiers bureautiques ou images.

L'entité Entrée générera les paquets d'information archivés (*AIP*) à partir des *SIP*, et les transmettra à l'entité Stockage. Elle s'assurera que ces paquets sont bien stockés (enregistrés) afin de transmettre des accusés de réception au service versant. Elle transmettra en parallèle à l'entité Gestion des données les données de conservation et de description des paquets archivés.

Enfin, toutes les entrées devront être journalisées afin de pouvoir tracer l'ensemble des transferts ou demandes de transferts effectués.

III.2.1.2 *Entité Stockage*

Le rôle de cette entité est essentiel. Elle permettra en effet de pérenniser la conservation des documents à travers le temps. Elle recevra donc les paquets d'information et gérera leur conservation physique sur les espaces de stockage dédiés.

Elle accusera réception de ces AIP afin d'informer l'entité Versement que le transfert s'est bien effectué. Cet accusé ne devrait être transmis qu'une fois l'écriture réellement effectuée, et non pas lors de la mise en file d'attente dans une mémoire tampon. En cas de coupure inattendue de la machine, cette mémoire serait en effet vidée, et l'information serait alors perdue.

Il est conseillé de diversifier les supports de stockage (*cf. Conseils généraux pour la pérennité des supports de stockage*

).

Si ce conseil est appliqué, il faudra alors que l'entité stockage gère, en fonction des fréquences de consultation estimées, le bon support de stockage pour enregistrer les AIP. En effet, en fonction du volume du paquet et de la fréquence de consultation, les supports choisis pourront varier.

En plus de stocker les AIP, l'entité Stockage pourra également les éliminer le cas échéant. Cette élimination devra pouvoir se faire automatiquement ou manuellement, et sera journalisée.

L'entité devra veiller à l'intégrité des informations archivées, en contrôlant automatiquement et fréquemment les données stockées. Afin de parer à ce genre de problème, il serait bon de prévoir une redondance du système, afin de pouvoir restaurer une information dont l'intégrité serait remise en cause. Elle devra également gérer les migrations de support en cas d'obsolescence technologique par exemple.

L'entité stockage gère également les moyens de sauvegarde du système (via l'entité Administration qui sera vu après).

Enfin, c'est elle qui transmettra les AIP à l'entité Accès lors des demandes de consultation.

III.2.1.3 Entité Gestion de données

L'entité Gestion de données traite des données descriptives des AIP mais également des données système. Comme l'entité Stockage, elle gère la conservation des informations, mais d'un point de vue logique. Elle englobe donc l'administration de la base de données, le maintien de son intégrité, de ses sauvegardes, la maintenance des schémas et des vues...

Elle fournit également les requêtes nécessaires à l'interrogation lors de la recherche de documents, et permet donc la génération de rapports. De ce fait, elle gère donc le lien entre les données physiques (entité Stockage) et les données descriptives qu'elle contient pour une archive.

Elle doit également rendre faisable la mise à jour de la base de données, qui sera lancée à partir de l'entité Administration.

III.2.1.4 Entité Administration

L'entité Administration inclut l'exploitation, la sécurité, le paramétrage, les migrations, le maintien en condition opérationnelle du système.

Dans la partie paramétrage, elle permet d'un point de vue fonctionnel de définir les accords de versement (formats autorisés, métadonnées descriptives obligatoires...), les acteurs (service versant, producteur, archives, administrateur technique, administrateur fonctionnel). Elle gère donc les comptes et profils utilisateurs du système.

Elle permet également d'effectuer les sauvegardes, les restaurations, les mises à jour, les migrations de format ou de support.

Elle permet de lancer des contrôles sur les données physiques, d'exécuter des requêtes d'utilisation, de visualiser les journaux de transaction...

C'est elle qui veille à la mise en place de sécurité d'accès au système, via l'authentification, les droits d'accès. Elle se basera de préférence sur le RGS.

En cas de redondance des informations, elle gérera également la réPLICATION d'un site vers l'autre.

L'entité administration devra également respecter le plus possible le RGAA et le RGI, afin de rendre interopérable le système, permettant alors sa réversibilité.

III.2.1.5 Entité Planification de la pérennisation

Cette entité assure la veille technologique liée à l'archivage, aussi bien pour les formats, les supports de stockage, mais également pour les nouvelles normes à mettre en place. Elle doit donc permettre de tester, de prototyper les futures évolutions du système.

Elle définit également la modélisation des paquets d'information AIP et SIP.

III.2.1.6 Entité Accès

L'entité Accès gère les demandes d'accès au système. Elle vérifie donc les droits d'accès des différents utilisateurs, mais est également là pour gérer les demandes de consultation et de restitution d'archives ou encore d'assistance sur le SAE.

Elle traite donc les réponses à transmettre aux usagers. En cas de validation d'une demande, elle permet d'exécuter les requêtes nécessaires à la recherche et la diffusion d'informations. Pour effectuer cette communication, elle génère des paquets d'information diffusés (*DIP*), qu'elle peut remettre en forme.

III.2.2 Synthèse des exigences

À partir de ces fonctionnalités au sens OAIS, nous pouvons synthétiser ces exigences sous forme de tableau, en y ajoutant quelques fonctionnalités plus techniques.

Les références d'exigence utilisent les préfixes suivants :

- FONC utilisé pour les exigences fonctionnelles.
- NFONC utilisé pour les exigences non fonctionnelles.

Tableau V : Liste des exigences du SAE

Identifiant	Libellé	Description
FONC010	Versement des archives	Les services versants (ou producteurs) devront pouvoir verser les documents à archiver dans ou vers le système avec les données descriptives
FONC020	Restitution d'archives	Il devra être possible de restituer des documents d'archives
FONC030	Élimination d'archives	Il devra être possible d'éliminer des documents d'archives
FONC040	Communicabilité des archives	Il devra être possible de gérer la communicabilité des documents d'archives
FONC050	Rechercher des archives	Il devra être possible d'effectuer des recherches sur les documents d'archives
FONC060	Consultation d'archives	Le système devra permettre de consulter des archives
FONC070	Création des accords de versement	Il devra être possible de créer des accords de versement d'archives (SEDA)
FONC080	Gestion des métadonnées	Le SAE devra permettre de définir les données descriptives
FONC090	Effectuer des demandes	Il devra être possible d'effectuer des demandes dans le SAE (versement, restitution, communication...)
FONC100	Validation / Rejet des demandes	Les différentes demandes (versement, restitution, communication...) pourront être acceptées ou refusées, soit de façon automatique (accord de versement non respecté), soit manuellement par certains utilisateurs
FONC110	Statistiques	Le système permettra d'effectuer des statistiques, de générer des rapports
FONC120	Notifications	Le SAE devra pouvoir notifier, alerter certains utilisateurs dans certains cas
FONC130	Contrôles	Le système devra pouvoir effectuer des contrôles (stockage, formats...) de façon automatique ou manuelle
FONC140	Gestion des utilisateurs	Le système permettra une gestion des comptes utilisateurs
FONC150	Gestion des droits	Le système permettra une gestion des droits utilisateur

Identifiant	Libellé	Description
FONC160	Accès sécurisé	L'accès au système se fera par authentification et les droits d'accès seront contrôlés
FONC170	Journalisation	Tous les mouvements, actions, connexions dans le SAE devront être journalisés et datés
FONC180	Gestion des formats	Le système devra pouvoir gérer des formats de documents autorisés ou pas
FONC190	Conversion de formats	Le système devra pouvoir convertir des formats
FONC200	Stockage des archives	Le système devra stocker les documents d'archives
FONC210	Conservation des archives	Le système doit conserver les documents archivés sur le moyen / long terme
FONC220	Maintenance du système	Le système sera maintenable par des administrateurs (mise à jour, migration des données, sauvegardes...)
FONC230	Migration des documents d'archives	Il devra être possible de migrer les documents d'archives vers un autre support de stockage
FONC240	Interconnexion avec systèmes	Il devra être possible d'interconnecter le SAE avec des systèmes externes (GED, autres SAE...)
FONC250	Intégrité des archives	Les documents archivés devront être non modifiables et conserver leur intégrité
FONC260	Confidentialité	Le système devra pouvoir gérer la confidentialité
FONC270	Réversibilité des archives	Il sera possible de quitter le système vers un autre système
FONC280	Possibilité de varier les supports de stockage	Le système devra supporter différents types de supports de stockage
NFONC010	Continuité de service	Le système devra être opérationnel et disponible
NFONC020	Lisibilité des archives	Les archives devront rester lisibles et compréhensibles sur le long terme
NFONC030	Évolutivité	Le système devra pouvoir être évolutif
NFONC040	Respect du standard d'échange	Le SAE devra respecter le standard SEDA
NFONC050	Respect des référentiels	Respect des référentiels RGAA, RGI et RGS
NFONC060	Respect modèle OAIS	Le SAE devra respecter le modèle OAIS
NFONC070	Respect norme	Le SAE devrait respecter la norme NF Z42-013 / ISO 146411
NFONC080	Respect de normes descriptives	Le SAE devra respecter des normes pour la description des données
NFONC090	Hébergement dans la collectivité	La solution d'archivage patrimonial devra être hébergée dans les locaux de la collectivité
NFONC100	Conservation sur plusieurs sites géographiques	Les archives devront être dupliquées sur différents sites géographiques

Ces exigences doivent permettre de couvrir les différents aspects nécessaires pour garantir la confiance dans le SAE et la pérennité des archives, à savoir les aspects concernant :

- l'intégrité
- la confidentialité
- la traçabilité
- la sécurité
- la compréhensibilité
- la lisibilité
- la réversibilité
- l'évolutivité

Précisons que dans le cadre de l'archivage patrimonial, la non-répudiation n'a pas réellement d'intérêt. Elle pourrait être implémentée avec des mécanismes tels que la signature électronique, mais quel serait l'intérêt de prouver dans plusieurs centaines d'années qu'un document a bien été signé par la personne qui le prétendait à l'époque. De plus, pour qu'une signature électronique soit valable, elle doit être certifiée par une autorité de certification. Or, rien ne prouve que les autorités existantes aujourd'hui existeront encore d'ici quelques années. Il faudrait alors, afin de conserver la valeur de la signature, conserver également tout le cheminement de la certification, ce qui n'aurait plus de sens dans de nombreuses années.

Toutefois, si le SAE doit également faire office de SAE intermédiaire, ce qui n'est pas le cas ici, il faudrait alors rendre cette exigence supplémentaire obligatoire.

III.3 Architecture type de la solution

La solution n'étant pas encore retenue, l'architecture ne l'est pas également. Toutefois, étant donné les fonctionnalités auxquelles doit répondre le futur système, nous pouvons déjà établir une architecture type qui sera proche de celle de la solution choisie.

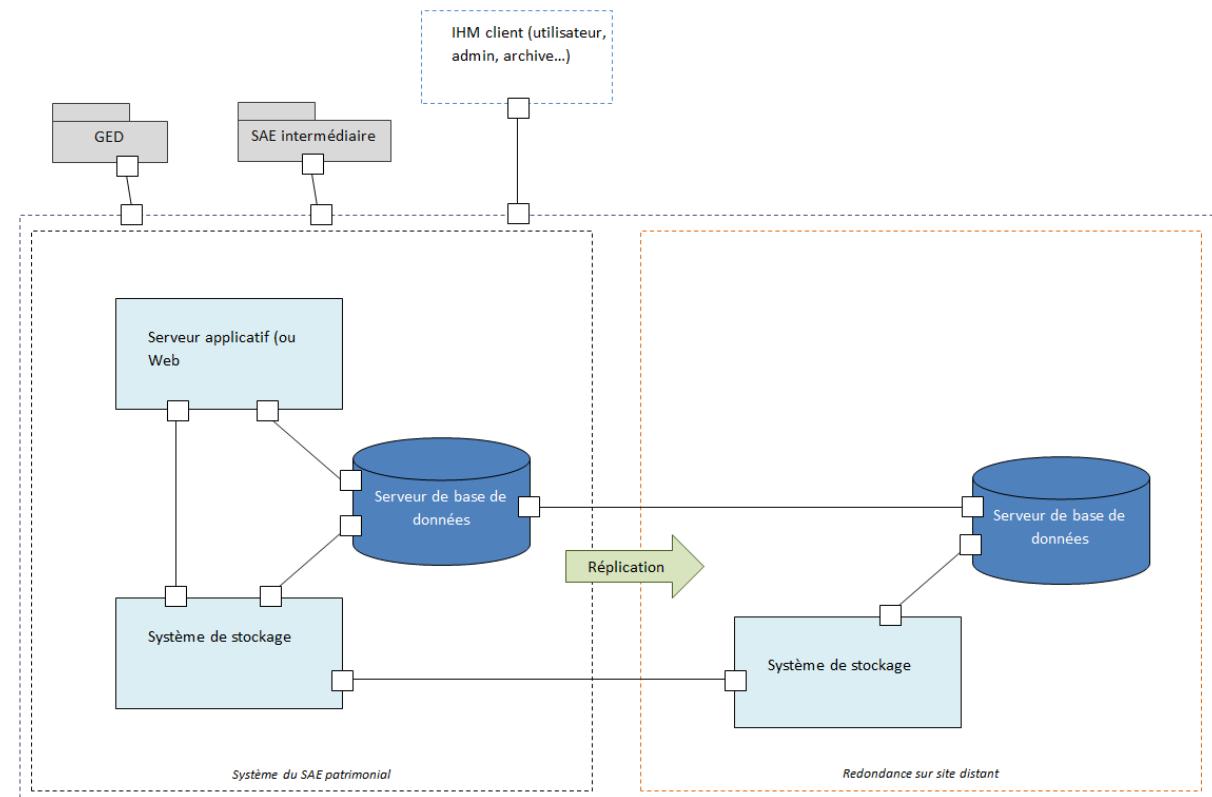


Figure 17: Architecture cible du futur système

Sur ce schéma, nous distinguons le système de SAE complet, qui est découpé en un système principal, et un système redondant, en cas de panne ou de problème sur le site maître.

Le système disposera de plusieurs interfaces. Au moins une interface IHM (**Interface Homme Machine**) permettant aux différents utilisateurs de se connecter au système. Il contiendra également des interfaces pour interconnecter le système avec les systèmes de GED et de SAE intermédiaire.

Sur le site principal, le SAE contiendra un serveur applicatif (ou serveur Web) qui accédera d'une part au serveur de base de données et d'autre part au système de stockage des documents. Il existera également un lien entre le serveur de base de données et le système de stockage.

Ces deux éléments seront répliqués sur un site distant. Sur le schéma, le serveur applicatif n'est pas répliqué. Il serait en effet plus facile de remonter ce serveur en cas de crash grâce aux sauvegardes et/ou s'il s'agit de serveurs virtuels. De plus, l'obligation pour un SAE réside dans l'intégrité des données et non dans la disponibilité continue du système.

III.4 Étude de marché

Nous avons décidé d'effectuer une petite étude de marché afin de présenter quelques produits existants sur le marché en matière d'archivage électronique. Il faut toutefois préciser qu'il existe d'autres solutions que celles présentées ici et que certains fournisseurs ont refusé de nous transmettre certaines informations techniques telles que le type de connecteurs utilisés pour l'interconnexion par exemple.

À noter que le CINES propose une étude de marché très complète datant de mai 2013, accessible à l'adresse suivante :

http://telechargement.girondenumerique.fr/databis/Doc_Archivage/Solutions_archivage%20%C3%A9lectronique.pdf

III.4.1 Solutions étudiées

À l'origine nous avions intégré plus de solutions pour cette étude, mais n'ayant pas réussi à récupérer suffisamment d'informations descriptives, nous ne les présentons pas ici.

Nous avons décidé de comparer les solutions suivantes pour cette étude :

- Arcsys



Le logiciel Arcsys est la solution d'archivage électronique proposée par le groupe Infotel. Elle propose l'archivage légal et l'archivage patrimonial, autant de manière externalisée (stockage externe au client) que de manière internalisée. Le groupe Infotel est également Tiers Archiveur agréé en France. Parmi ses partenaires, nous pouvons citer Axa, Groupama, EDF, ERDF et Bouygues Télécom.

- As@lae



As@lae est le SAE proposé par l'Adullact. Il s'agit d'un système totalement Open Source, traitant aussi bien de l'archivage intermédiaire que l'archivage pérenne. As@lae est une plateforme qui s'installe uniquement chez les clients. Cette solution est utilisée par Mégalis pour l'archivage probant des collectivités en Bretagne, mais également sur le SAE mutualisé entre Bordeaux Métropole, la Ville de Bordeaux et le Département Gironde.

- CDC Arkhinéo



CDC Arkhinéo, filiale de la Caisse des Dépôts, gère également les archives légales et patrimoniales. Par contre, la solution n'est disponible qu'en mode SAAS (*Software As A*

Service). Il s'agit donc d'une solution de type Cloud, dont les datacenters sont toutefois installés en France. CDC Arkhinéo est aussi Tiers Archiveur agréé par le Ministère de la Culture. Parmi ces partenaires, nous pouvons citer les groupes Carrefour, Leclerc, BNP Paribas, le Crédit Agricole, Dassault Aviation.

- Maarch RM



Tout comme As@iae, Maarch RM est un logiciel conçu sur des briques Open Source. C'est la partie SAE d'une suite logicielle contenant également un système de gestion électronique de documents (*GED*) et d'un système de gestion électronique de courriers (*GEC*). Beaucoup de collectivités locales en France utilisent la solution Maarch RM.

- Mnesys Archives



Mnesys Archives est édité par la société Naoned. C'est un logiciel qui traite de l'archivage physique et électronique. Le Ministère des Affaires Étrangères, le Conseil Général de la Manche, la Ville de Quimper et l'Assistance Publique des Hôpitaux de Paris sont utilisateurs de cette solution.

- Spark Archives



Spark Archives est le SAE développé par le groupe Klee. Cette solution gère également les archives papiers et les archives numériques. Plusieurs ministères tels que le Ministère de l'Écologie ou le Ministère de l'Économie et des Finances, du Développement Durable et de l'Énergie, le groupe Total, la Caisse des Dépôts et des Consignations ou encore la SNCF sont utilisateurs de Spark Archives.

III.4.2 Comparatif des différentes solutions étudiées

Le tableau ci-après présente certaines des fonctionnalités offertes par ces solutions. Toutefois, pour certaines fonctionnalités, nous ne pouvons pas indiquer si les solutions en disposent, du fait du manque d'information dont nous disposons.

L'ensemble de ces solutions disposent des fonctionnalités de journalisation, d'horodatage des documents, de réversibilité, de gestion des droits utilisateurs, ainsi que des diverses étapes pour l'archivage, à savoir le versement, la recherche, la communication, la restitution, l'élimination.

Tableau VI : Comparatif des fonctionnalités des solutions SAE

Solutions Fonctionnalités		Arcsys	As@lae	CDC Arkhinéo	Maarch RM	Mnesys	Spark Archives
Type d'import		Interface Web + web services	Interface Web + web services	Web Services	Interface Web + Batch	Interface Web	Interface Web
Connecteurs / API		Web services et autres API ¹ (?)	RESTful ² - SOAP ³	RESTful	Pas actuellement	CMIS ⁴ – Interface avec as@lae	CMIS - IMAP
Hébergement	Externe (SAAS)	OUI	NON	OUI	NON	OUI	OUI
	Interne	OUI	OUI	NON	OUI	OUI	?
Gestion des formats		Certains formats	OUI	?	OUI (non paramétrable pour l'instant)	OUI	OUI
Gestion des supports de stockage		Disque dur et Bande magnétique	OUI	OUI	?	OUI	?
Conversion des formats		Pas actuellement	OUI	?	Pas actuellement	OUI	OUI
Migration des volumes		?	OUI	OUI	?	OUI	?
Normes	OAIS	OUI	OUI	?	OUI	OUI	OUI
	ISO-15489	OUI	OUI	OUI	OUI	OUI	OUI
	NF Z42-013	OUI	OUI	OUI	OUI	OUI	OUI
	SEDA	OUI	OUI	OUI	OUI	OUI	OUI
Répartition géographique		OUI	OUI	OUI	OUI	OUI	OUI
Open Source		NON	OUI	NON	OUI	NON	NON
Tarification		De 70 000 à 250 000 €		?	?	40 000 à 50 000 €	En fonction du volume

¹ API : Application Programming Interface ou Interface de programmation² RESTful : par abus de langage, désigne un service web basé sur le style d'architecture REST (Representational State Transfer)³ SOAP : Simple Object Access Protocol⁴ CMIS : Content Management Interoperability Services

III.4.3 Choix de la solution retenue pour la maquette

Les solutions hébergées en mode Cloud, directement chez le prestataire, apporteraient un confort important. En effet diverses tâches telles que la disponibilité, l'intégrité, la pérennité... incomberaient à l'hébergeur et non à notre service informatique.

De même, les coûts liés à la consommation électrique, la climatisation, l'administration et la maintenance des serveurs et des espaces de stockage ne seraient pas à la charge de la collectivité.

Toutefois, étant donné l'état actuel de la législation concernant les archives publiques définitives, ce genre de solution n'est pas envisageable.

Dans le cadre de notre prototype, la solution As@lae semble la plus indiquée. Outre le fait qu'elle soit Open Source et gratuite, elle répond bien aux exigences spécifiées auparavant. Il faut toutefois nuancer le terme « *gratuit* », car les éditeurs, tels que l'Adullact, proposent l'intégration et la formation à ces solutions, qui sont évidemment payantes. As@lae est de plus le SAE intermédiaire proposé par Mégalis et retenu pour la Ville de Rennes et Rennes Métropole, permettant ainsi une meilleure interopérabilité entre les deux systèmes.

Toutefois, l'autre solution Open Source étudiée, Maarch RM semble également être très intéressante. Elle ne dispose pas encore de toutes les fonctionnalités dont dispose As@lae, mais étant une solution très récente, ces fonctionnalités sont incluses dans la Road Map et devraient donc faire partie de la solution prochainement.

Lorsque la Ville de Rennes et Rennes Métropole décideront de lancer le projet de mise en place d'un SAE patrimonial, un cahier des charges précis sera rédigé. Le choix de la future solution dépendra des réponses au futur dossier de consultation qui sera alors publié.

III.4.4 Autres solutions

En plus de ces solutions étudiées dans le paragraphe précédent, il semble important de mentionner une autre approche. Dans le cadre de la modernisation de l'action publique, l'État a en effet lancé un appel d'offre pour réalisation d'un SAE Open Source.

Il s'agit du programme VITAM (*Valeurs Immatérielles Transférées aux Archives pour Mémoire*), issu de la collaboration des Ministères de la Culture et de la Communication, de la Défense et des Affaires Étrangères de du Développement International [27]¹.

¹ Secrétariat Général pour la Modernisation de l'Action Publique : Vitam : Vers un socle d'archivage électronique commun à toute l'administration

Ce futur SAE répondra aux exigences en matière de normes (ISO 15489, ISO 14721, NF Z42-013) et devraient permettre, grâce à divers connecteurs, d'interconnecter les différents logiciels métiers.

Lorsque son développement sera achevé, cette brique offrira un socle commun aux administrations de l'État. Dans un deuxième temps, la solution d'archivage électronique devrait être étendue aux collectivités territoriales. Il sera normalement possible d'installer cette brique en interne, et d'effectuer des développements spécifiques dessus.

Par conséquent, VITAM pourra se révéler intéressant pour Rennes Métropole lorsque son développement sera achevé. C'est donc une piste à suivre d'ici les prochaines années.

La spécification ayant été faite et le choix s'étant porté sur la solution As@iae, nous pouvons maintenant passer à la mise en œuvre du prototype.

IV Réalisation et mise en œuvre du prototype

Comme nous l'avons précisé, le choix effectué pour la réalisation de la maquette est la solution As@lae de l'Adullact. Cette partie entend présenter cette mise en œuvre, ainsi que la réalisation d'un connecteur avec la GED Alfresco interne à Rennes Métropole. Il était prévu initialement d'effectuer un envoi de fichier depuis le SAE d'archivage intermédiaire As@lae de Mégalis, mais la plateforme n'étant toujours pas mise en place, ce test n'a pas pu être réalisé.

Nous présenterons entre autre l'installation effectuée, les tests réalisés sur certaines des fonctionnalités de la solution, comme le dépôt manuel d'archives via la plateforme, le dépôt via les connecteurs entre la solution As@lae et la GED ou encore la création de profils d'archivage. Enfin, nous proposerons une évaluation de ce prototype.

IV.1 Installation de la solution

Dans le cadre de cette maquette, afin de me permettre d'utiliser la maquette en dehors du réseau de Rennes Métropole, il a été décidé d'installer les différents composants sur une seule machine physique avec des machines virtuelles. Dans le cas d'une mise en œuvre réelle, l'architecture retenue serait bien évidemment différente.

Le poste physique ayant servi à l'installation tourne sous Windows 7 64 bits. À partir de l'outil Oracle Virtual Box, un serveur virtuel Ubuntu 12.04 LTS Server a été créé sur ce poste. C'est sur cette machine virtuelle qu'ont été installés les différents composants d'As@lae.

IV.1.1 Installation des composants

Nous présentons ci-dessous l'installation des prérequis et de certains des composants nécessaires pour le fonctionnement d'As@lae. Il faudra se référer à la documentation d'installation de l'Adullact pour avoir la liste complète des commandes à exécuter pour l'installation de tous les divers composants du SAE.

IV.1.1.1 Prérequis

As@lae nécessite plusieurs composants pour son fonctionnement ou son installation, tels qu'un antivirus, la possibilité de se connecter à un serveur de temps ou encore à un serveur de messagerie (pour l'envoi des notifications).

- Installation des prérequis (**postfix** pour les mails, **ntp** pour se connecter à un serveur de temps, **clamav** comme antivirus...) :

```
root@testsaejm# aptitude install ssh vim postfix rsync unzip bsd-mailx ntp
python-software-properties
root@testsaejm# aptitude install clamav clamav-freshclam
```

- Après cette étape, il faudra paramétrer les fichiers de configuration pour **postfix** en définissant le relai SMTP (*Simple Mail Transfer Protocol*), ainsi que les serveurs de temps (internes au réseau ou externes selon les cas). Il faut également paramétrer, si besoin, le proxy pour l'accès internet.
- Le JRE (*Java Runtime Environment*) version 6 de Java est également nécessaire, il faut donc le télécharger et l'installer.
- As@lae fonctionne sur un serveur Web **Apache**, il faut par conséquent installer les prérequis du serveur Web :

```
root@testsaejm# aptitude install curl php5 php-soap php5-pgsql php5-xsl
php5-curl php-pear
root@testsaejm# aptitude install apache2
root@testsaejm# aptitude install XML_RPC2
```

- Configurer directement le fichier de configuration de **php.ini** en positionnant les valeurs suivantes
 - post_max_size = 100M
 - upload_max_filesize = 100M
 - expose_php = Off
 - date.timezone = « Europe/Paris »
 - max_execution_time = 3600
 - memory_limit = 1024M # ou 512

IV.1.1.2 Récupération et installation du serveur Web d'as@lae

As@lae fonctionne sur un serveur Web **Apache**, qu'il faudra paramétrer afin de pouvoir se connecter au serveur en https. On pourra toutefois, si on le souhaite, utiliser as@lae directement en http.

- Récupérer les sources d'As@lae (écrit en **cake-php**) :

```
root@testsaejm:/var/www# wget
http://adullact.net/frs/download.php/7294/asalae.1.5-cakephp1.2.10.zip
```

- Décompresser les sources d'As@lae.
- Activer les modules d'apache et création d'un répertoire pour héberger les certificats (notre serveur virtuel **Apache** se nommera asalae) :

```
root@testsaejm# a2enmod rewrite ssl
root@testsaejm# mkdir -p /etc/apache2/ssl/asalae
```

- Générer un certificat (ici en test uniquement) pour le vhost https

```
root@testsaejm# sudo openssl req-x509 -nodes -days 365 -newkey rsa:2048 -keyout /etc/apache2/ssl/asalae/asalae.key -out /etc/apache2/ssl/asalae/asalae.crt
```

```
Generating a 2048 bit RSA private key
.....+++
.....+++
writing new private key to '/etc/apache2/ssl/asalae/asalae.key'
-----
You are about to be asked to enter information that will be incorporated
into your certificate request.
What you are about to enter is what is called a Distinguished Name or a DN.
There are quite a few fields but you can leave some blank
For some fields there will be a default value,
If you enter '.', the field will be left blank.
-----
Country Name (2 letter code) [AU]:FR
State or Province Name (full name) [Some-State]:Bretagne
Locality Name (eg, city) []:Rennes
Organization Name (eg, company) [Internet Widgits Pty Ltd]:RM
Organizational Unit Name (eg, section) []:DSI
Common Name (e.g. server FQDN or YOUR name) []:testsaejm
Email Address []:jm@test.local
```

Figure 18 : Génération d'un certificat

- Il faut ensuite copier et paramétriser les vhosts (Virtual Hosts : Serveurs virtuels au sein du serveur Apache) en http et/ou en https :

```
root@testsaejm# cp /var/www/asalae/app/config/apache2/* /etc/apache2/sites-available/asalae.ssl
root@testsaejm# vim /etc/apache2/sites-available/asalae.ssl
```

- Extrait du fichier de configuration du vhost (**asalae.ssl**) pour accéder en https à as@lae :

```
<Virtual Host *:443>
  ServerName      asalae.mutu.local
  ServerAlias     asalae
  DocumentRoot   /var/www/asalae
  AddDefaultCharset    utf-8
  SSLEngine       on
  SSLCertificateFile /etc/apache2/ssl/asalae/asalae.crt
  SSLCertificateKeyFile /etc/apache2/ssl/asalae/asalae.key
[...]
```

- Il faut ensuite activer le vhost créé, vérifier sa syntaxe et redémarrer le serveur apache :

```
root@testsaejm# a2ensite asalae.ssl
root@testsaejm# apache2ctl -t
root@testsaejm# service apache2 restart
```

IV.1.1.3 Outil de validation des formats

As@lae utilise l'outil **FormatValidator** du CINES pour la validation des fichiers d'archives, et l'utilitaire Jing pour la validation des fichiers au format RelaxNG (**Regular Language for XML Next Generation**). Le format RelaxNG permet de générer les profils d'archives. Il s'agit d'un langage de description étant une alternative au schéma XML.

- Il faut au préalable installer le compilateur GCC pour les applications C et C++. Il faudra ensuite lancer un script fournit dans les sources d'as@lae qui ira télécharger et installer **FormatValidator** depuis le site du CINES :

```
root@testsaejm# aptitude install make gcc g++ ia32-libs
root@testsaejm# /var/www/asalae/app/config/cines/install-cines-v1.3.sh
root@testsaejm# aptitude purge make gcc g++
```

- **FormatValidator** sera installé dans le répertoire /opt (chemin configuré dans le script d'installation). Il faudra ensuite paramétrier éventuellement dans /opt/formatValidator2 le fichier **launch_asalae.sh** afin de définir le JAVA_HOME et son emplacement.
- Il est également possible d'utiliser la validation de format de fichier en mode hors-ligne, car par défaut, cette validation demande un accès internet. Dans ce cas, il faut éditer le fichier /opt/formatValidator2/etc/jhove.conf et remplacer comme dans la copie d'écran suivante :

```
<jhoveConfig version="1.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://hul.harvard.edu/ois/xml/ns/jhove/jhoveConfig"
  xsi:schemaLocation="http://hul.harvard.edu/ois/xml/ns/jhove/jhoveConfig
    http://asalae/app/webroot/files/jhoveConfig.xsd">
```

Figure 19 : Paramétrage du fichier jhove.conf

- Pour l'installation de **Jing**, il suffit de télécharger les sources, de les décompresser et de créer un lien symbolique vers un répertoire nommé Jing :

```
root@testsaejm# wget https://jing-trang.googlecode.com/files/jing-
2009111.zip
root@testsaejm# ln -s jing-20091111 /opt/jing
```

IV.1.1.4 Outil de conversion des fichiers

Pour la conversion de fichier, il est possible d'installer **Cludoooo** ou **Unoconv**. Ces deux utilitaires utilisent Libre Office ou Open Office pour la conversion de fichier.

As@lae utilise également l'utilitaire **ODFGedo00** permettant la génération et la fusion de fichier PDF.

Nous ne sommes pas parvenus à faire fonctionner ces deux composants. Malgré la documentation d'installation et les forums sur Ubuntu, les deux outils se lancent mais ne convertissent aucun document.

IV.1.1.5 Bases de données

As@lae utilise également le SGBDR **Postgresql**, en version 9.1. Dans le répertoire d'as@lae est fourni un ensemble de scripts sql permettant la génération de la base de données avec le contenu minimal.

- Installation de **postgresql** :

```
root@testsaejm# aptitude install postgresql-9.1 postgresql-client-9.1
postgresql-contrib-9.1 postgresql-doc-9.1 postgresql-plperl-9.1 postgresql-
plpython-9.1 postgresql-pltcl-9.1
```

- Il faut ensuite paramétrer l'autorisation d'accès à la future base de données. Nous autorisons en local, le user 'asalae' à accéder en 'MD5' à la base de données 'asalae' en éditant le fichier **pg_hba.conf** de **postgresql** :

```
#TYPE      DATABASE        USER        ADDRESS METHOD
local    asalae          asalae      md5
```

- Une fois ce fichier configuré, il faut créer la base de données et l'utilisateur. Pour se faire, se connecter en *psql* sur le serveur, puis

```
postgres=# CREATE USER asalae;
postgres=# CREATE DATABASE asalae OWNER asalae;
postgres=# ALTER USER asalae WITH ENCRYPTED PASSWORD XXXXXXXXX;
```

- Il faut ensuite se connecter à cette base de données que l'on vient de créer avec l'utilisateur asalae et lancer le script de création **asalae_postgres_1.5.sql** présent dans le répertoire /var/www/asalae/app/config/postgresql/. Ce script va créer les tables et données de configuration d'as@lae.

IV.1.1.6 Horodatage

Il reste à installer un autre composant important du SAE, l'outil permettant d'horodater les fichiers. Il s'agit de **LiberHorodatage**, projet également développé par l'Adullact. Cet outil est également à installer en tant que vhost **Apache**. L'Adullact fournit également une documentation d'installation de ce composant.

IV.1.1.7 Fichiers de configuration

Il faut enfin effectuer quelques paramétrages afin de rendre as@lae fonctionnel.

Tout d'abord, il faudra mettre le compte d'Apache (www-data) propriétaire des répertoires contenant le site web du SAE.

Il faut également créer les répertoires qui contiendront les échanges de documents entre as@lae et les utilisateurs. Par défaut, les répertoires sont construits comme ci-dessous :

```
+---data
| +---echanges
| | +---entree
| | +---sortie
| +---messages
| +---archives
```

- Enfin, pour terminer, il faudra personnaliser les fichiers de configuration **database.php**, **types_connecteurs.ini.php** et **asalae.ini.php** présents dans le répertoire /var/www/asalae/app/config.

IV.1.2 Architecture d'as@lae

Même si tous les composants ont été installés sur la même machine physique, l'architecture d'as@lae peut être décomposée de la façon suivante :

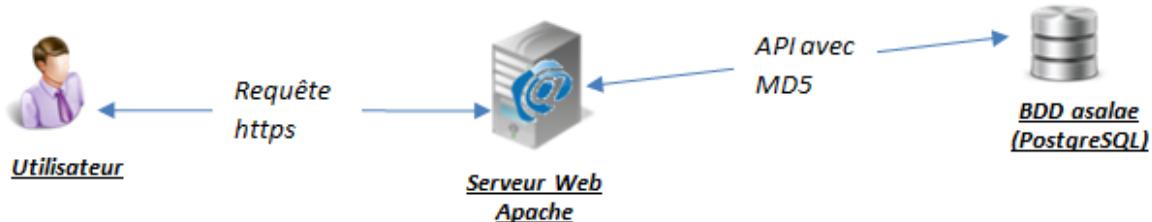


Figure 20 : Architecture d'As@lae

IV.2 Paramétrages as@lae

Notre maquette de SAE est maintenant installée. Il restera aux administrateurs fonctionnels de paramétrer certaines fonctionnalités avant d'effectuer les tests.

IV.2.1 Accès à l'application

L'application est accessible à l'adresse <https://asalae/>

Nous arrivons bien sur la page d'authentification :



Adullact Système Archivage Electronique

Entrez votre identifiant et votre mot de passe pour continuer.

Identifiant de connexion

Mot de passe

as@lae - V1.5 - © 2010-2014 ADULLACT

Figure 21 : Écran de connexion à l'application As@lae

Le login/mot de passe par défaut est admin/admin, qu'il faudra bien entendu modifier.

Une fois connecté, on arrive sur la page d'accueil d'as@lae :

[Accueil](#) [Transferts](#) [Entrées](#) [Communications](#) [Eliminations](#) [Restitutions](#) [Administration](#) [Administration SEDA](#) [Administration technique](#)

▽ Raccourcis vers les fonctions d'as@lae

Transferts

- [Transferts pré-formatés](#)
- [Transferts par formulaire](#)
- [Transferts non conformes](#)
- [Mes transferts à traiter](#)
- [Tous les transferts à traiter](#)
- [Tous les transferts](#)

Entrées

- [Registre des entrées](#)
- [Navigation par niveau](#)
- [Recherches](#)
- [Transferts archivage définitif](#)
- [Op. sur les entrées](#)

Communications

- [Demandes de communication](#)
- [Validation des demandes de communication](#)

Figure 22 : Page d'accueil d'As@lae

IV.2.2 Paramétrages fonctionnels

C'est maintenant que nous allons pouvoir configurer les comptes utilisateurs avec les droits d'accès, les acteurs d'archives, les types de connecteurs utilisés...

IV.2.2.1 Onglet Administration SEDA

C'est à partir de cet écran que vont être définis les acteurs SEDA tels que les services producteurs, versants, contrôleurs ou encore demandeurs. C'est également dans cette partie que nous allons créer les profils de versement.

IV.2.2.1.1 Acteurs SEDA

Nous créons ici un acteur SEDA qui pourra produire, verser et demander des archives. C'est cet acteur SEDA qui servira essentiellement dans nos tests de transfert d'archives.

Edition de l'acteur SEDA : Service versant DS1

Informations principales Traitements Paramétrage SEDA v0.2 Paramétrage SEDA v1.0

Nom
Service versant DS1

Identifiant unique
FRAD001

Rattaché à

Privé : l'acteur n'entre pas en jeu dans les échanges. Il est simplement déclaratif et peut servir par exemple à l'élaboration de la hiérarchie des services.

Type d'acteur SEDA
Service de contrôle
Service versant **Service producteur**
Service demandeur

Type d'entité
Collectivité

Figure 23 : Crédit d'un acteur SEDA sous As@iae

IV.2.2.1.2 Profil d'archives

Il convient maintenant de créer un profil d'archives. Il faut qu'on indique quels champs informations (ou métadonnées) pourront ou devront être présents dans le bordereau de transfert au format SEDA.

Un profil d'archive se base sur un fichier XML ou RelaxNG (qui est un langage de description de document XML).

Le logiciel **Agape** (*Application de Génération Automatisée de Profils Electronique SEDA*) est une application libre permettant de faciliter la génération de fichier au format RelaxNG.

Il s'agit d'un fichier .jar (application Java) qui ne nécessite aucune installation et est téléchargeable sur la forge de l'Adullact.

Il permet de définir les champs qui seront autorisés, obligatoires et facultatifs dans le bordereau de transfert SEDA.

Nous générerons donc un fichier de RelaxNG via l'outil **Agape** avant de pouvoir créer notre profil d'archives sous as@iae. Le début du fichier généré est visible en *Annexe 3 : Fichier RelaxNG pour profil d'archives*.

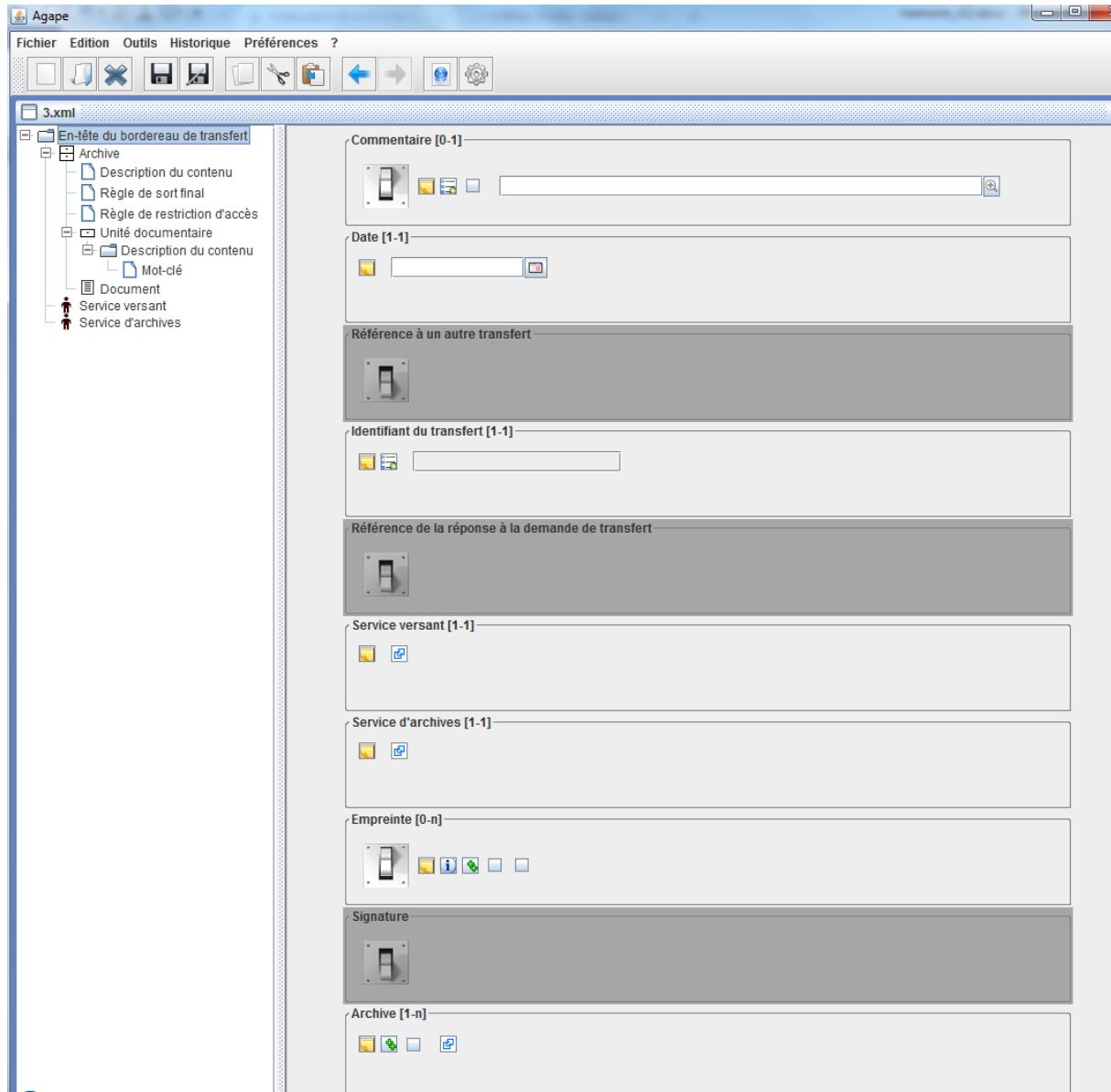


Figure 24 : Génération d'un fichier RelaxNG avec l'outil Agape

Après avoir défini les champs autorisés ou non, il suffit de générer le fichier .rng.

De retour sur l'interface d'As@iae, nous pouvons maintenant créer notre profil et y associer le fichier RelaxNG créé.

Nouveau profil

Informations principales

Attributs SEDA

Nom

Identifiant unique du profil : correspond à ArchivalProfile

Description

Fichier(s) de validation : XML Schema (*.xsd) et RelaxNG (*.rng)

+ Ajouter

seda_v1_schema.rng	11.17 KB
--------------------	----------

Figure 25 : Crédit d'un profil d'archivage sous As@lae

Notre profil est maintenant créé. Les futurs transferts d'archives utilisant ce profil devront respecter le fichier .rng pour la génération du bordereau de transfert.

IV.2.2.1.3 Accord de versement

Après avoir créé un profil d'archives, il faut créer un accord de versement auquel il sera associé. Il sera valable pour le service de contrôle (service des archives) et un ou plusieurs services versants. C'est dans cet accord que seront définis les formats de fichiers autorisés pour le dépôt d'archive.

Nouvel accord de versement

Informations principales Contrôles Alertes par courriel Attributs SEDA

Nom

Accord_versement_DSI

Identifiant : correspond à 'ArchivalAgreement'

ACC_VERS_DSI01

Description

Accord de versement entre le service versant DSI et les archives

Service d'archives

Service d'archives - FRAD000 ▾

Service(s) versant

Service versant Marché - FRAD003
Service versant DSI - FRAD001

Profil d'archive

test_profil_archivage ▾

Figure 26 : Crédit d'un accord de versement sous As@iae

Nous autorisons les formats XML et PDF en version 1.4, ce qui correspond à du PDF/A.

Nouvel accord de versement

Informations principales Contrôles Alertes par courriel Attributs SEDA

Signature du bordereau de transfert

Présence obligatoire de la signature. Note : le contrôle se fait sur la présence mais non sur la validité de la signature

Pièces jointes

Liste des formats autorisés

* fmt/18 - Portable Document Format (Version=1.4) * fmt/101 - Extensible Markup Language (Version=1.0)

Figure 27 : Définition des formats de fichiers autorisés pour un accord de versement sous As@iae

Après la création de cet accord, il nous faut encore créer un contrat lié à cet accord avec un service producteur. Ce contrat est juste un accord implicite entre le service versant, ici la DSI et le service producteur, la DSI également dans notre test.

IV.2.2.2 Onglet Administration Technique

L'onglet 'Administration Technique' permet de configurer les volumes de stockage, de créer

des mots-clés, de visionner les journaux de transaction, de planifier des tâches ou encore les types de connecteurs (mails, autre système As@lae) qui seront utilisés par notre SAE.

IV.2.2.3 Onglet Administration

L'onglet '*Administration*' permet le paramétrage de la ou des collectivités utilisatrices du logiciel, des utilisateurs, des droits d'accès, des services utilisateurs, des circuits de traitement.

Nous créons tout d'abord un service DSI (service utilisateur et non acteur SEDA comme fait précédemment) que nous relions à notre acteur SEDA, le service versant DSI.

Ensuite nous créons ici deux utilisateurs spécifiques :

- un utilisateur '*dsi*' qui peut effectuer des transferts, des demandes de communication, de restitution et d'élimination à partir de l'interface as@lae, qui est associé au service utilisateur DSI qui vient d'être créé (et par conséquent également à l'acteur SEDA 'service versant DSI').
- un utilisateur '*adminrest*' qui peut uniquement utiliser les web services RESTful d'as@lae.

IV.3 Tests depuis interfaces as@lae

Nous allons maintenant pouvoir tester réellement différentes fonctionnalités du SAE. Nous allons tout d'abord tester le transfert d'archives conformes et non conformes via la plateforme As@lae. Pour cela, nous nous connectons à l'interface avec l'utilisateur *dsi* précédemment créé.

Les droits sont bien définis, nous n'avons accès qu'aux modules auxquels nous lui avons donnés les droits.

The screenshot shows the As@lae homepage with a navigation bar at the top. On the right side, there is a user profile section for 'Rennes Métropole' with a dropdown menu for 'Service versant DSI' and a link to 'dsi dsi - [déconnexion]'. Below the navigation bar, there are five main menu items: Accueil, Transferts, Communications, Eliminations, and Restitutions. Under each of these, there is a list of sub-links. The 'Transferts' section has a single link: 'Transferts par formulaire'. The 'Communications' section has a single link: 'Demandes de communication'. The 'Eliminations' section has a single link: 'Demandes d'élimination'. The 'Restitutions' section has three links: 'Demandes de restitution', 'Restitutions', and 'Acquittement des restitutions'. A sidebar on the left contains a link 'Raccourcis vers les fonctions d'as@lae' which lists the same categories and their sub-links.

Figure 28 : Page d'accueil d'As@lae avec des droits restreints

IV.3.1 Versement fichier PDF version 1.4 et XML

Dans ce premier test, nous allons verser via l'interface Web d'As@lae un fichier PDF en version 1.4 joint avec un fichier XML contenant des métadonnées descriptives. Pour cela, nous avons généré le fichier PDF en question à partir du présent document (avec une version non terminée de celui-ci). Le fichier XML correspond à l'exemple des métadonnées METS, accessible en *Annexe 2 : Exemple METS*

Les écrans suivants montrent les étapes principales lors d'un versement manuel depuis l'interface :

Ajout d'un transfert en mode interactif

The screenshot shows the 'Entête du transfert' (Transfer Header) section of the interface. On the left, there is a tree view with nodes: 'Entête du transfert', 'Service versant' (selected), 'Service d'archives' (disabled), and 'Archives (0)'. On the right, there are several input fields:

- Commentaire:** 'premier transfert de la DSI vers SAE'
- Date du bordereau:** '20 - 08 - 2015' with a calendar icon.
- Identifiant du transfert:** A checkbox is checked with the label 'Calculé lors du verrouillage'.
- Référence à un autre transfert:** An empty input field with a help icon.
- Référence à la réponse à la demande de transfert:** An empty input field with a help icon.
- Valider:** A button at the bottom.

Figure 29 : Versement manuel d'archive depuis As@lae : entête du transfert

Nous définissons ensuite le service versant (le service DSI), ainsi que le service d'archives qui contrôlera le versement.

Après cette étape, il faut ajouter la notion d'archive, en précisant son nom, ainsi que l'accord de versement et le profil d'archivage utilisés.

The screenshot shows the 'Nom, intitulé de l'archive' (Name, title of the archive) section of the interface. On the left, there is a tree view with nodes: 'Entête du transfert', 'Service versant' (selected), 'Service d'archives' (disabled), and 'Archives (0)'. On the right, there are several input fields:

- Nom, intitulé de l'archive:** 'Version actuelle du mémoire JM'
- Niveau de description:** 'item - Item'
- Langue des descriptions:** 'fr - français'
- Accord de versement (convention d'archivage):** 'Accord_versement_DSI - Accord de versement entre le service versant DSI et les archives'
- Profil d'archive:** 'test_profil_archivage -'
- Niveau de service:** An empty input field with a help icon.
- Identifiant d'archive fourni par le service versant:** An empty input field with a help icon.
- Identifiant de l'archive attribué par le service d'archives:** An empty input field with a help icon.
- Valider** and **Annuler** buttons at the bottom.

Figure 30 : Versement manuel, intitulé de l'archive et choix du profil d'archivage et accord de versement

Avant d'ajouter les pièces jointes (les documents d'archives), nous définissons le sort final et la communicabilité des archives dans les écrans prévus. Ici, nous positionnons la conservation comme sort final et précisons que les documents sont communicables sans délai.

Nous ajoutons ensuite nos deux documents, tout d'abord le document PDF, en précisant qu'il s'agit de « contenu de données », alors que le document XML METS sera de type « information de pérennisation ».

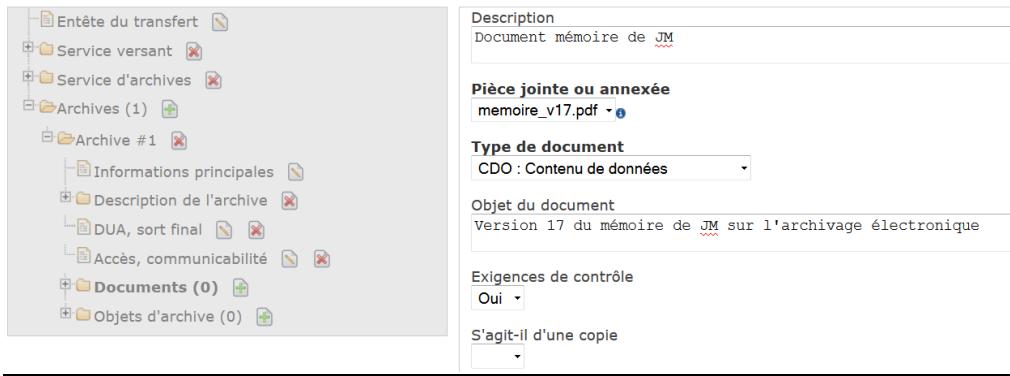


Figure 31 : Ajout du document PDF en tant que contenu de données

Description	Métadonnées METS associées au mémoire version 17 de JM
Pièce jointe ou annexée	mets_mem_pdf.xml
Type de document	PDI : Information de pérennisation
Objet du document	Métadonnées descriptives, de provenance

Figure 32 : Ajout du fichier de métadonnées METS

Une fois le versement créé, on peut visualiser les pièces jointes ajoutées et les contrôles effectués.

Vue détaillée du transfert : #A_CALCULER_LORS_DU_VERROUILLAGE#

[Transfert](#) [Pièces jointes](#) [Info antivirus](#)

Pièce jointe 'memoire_v17.pdf'

Fichier	memoire_v17.pdf [Télécharger]
Taille	2.8 Mo (2 924 615 Octets)
Validation	effectuée par l'outil de validation 'CINES' le 28-08-2015 à 10:27:41
Résultat validation	bien formé : Oui, valide : Oui, archivable : Oui
Info. issues de la validation	format : PDF, version : 1.4

Pièce jointe 'mets_mem_pdf.xml'

Fichier	mets_mem_pdf.xml [Télécharger]
Taille	4.1 Ko (4 187 Octets)
Validation	effectuée par l'outil de validation 'CINES' le 28-08-2015 à 10:27:57
Résultat validation	bien formé : Oui, valide : Oui, archivable : Oui
Info. issues de la validation	format : XML, version : 1.0

Figure 33 : Visualisation des pièces jointes ajoutées lors de la création du versement

On constate que le système a généré des fichiers pour horodater les pièces jointes des archives (fichiers .tsa). Le système a également vérifié la conformité des formats des fichiers, ainsi que la non-présence de virus.

L'utilisateur DSI peut maintenant verrouiller le transfert et le transmettre pour contrôle au service des archives. Le bordereau SEDA de transfert a alors été généré par As@lae, on peut le voir en *Annexe 4 : Bordereau de transfert SEDA*

Ensuite, en se connectant avec l'utilisateur des archives, acteur SEDA pouvant effectuer le contrôle sur le versement, le contrôle du transfert est effectué et accepté par les archives.

Validation du transfert : AT_32

Traitement **Transfert** Historique des traitements

Choisissez le traitement à appliquer

Accepter

Commentaire du traitement (interne)
ok

Commentaire du message de notification d'acceptation ou d'avis d'anomalie
votre fichier est accepté

Accepter Annuler

Figure 34 : Validation du versement par le service des Archives

Une fois cette étape passée, les documents archivés font partie du SAE. On peut les voir dans les registres des entrées.

Accueil	Transferts	Entrées	Communications	Eliminations	Restitutions	Administratio
Registre des entrées		Navigation par niveau	Recherches	Transferts archivage définitif		Op. :

Liste des entrées

Trier par : [Cote](#) [Date création](#) [Taille](#) [Producteur](#) [Versant](#) [Service d'archives](#)

4 éléments affichés sur un total de 4

Tout cocher, tout décocher

<input type="checkbox"/>	4 : Version actuelle du mémoire JM
	Sort final : Conserver . DUA : 2 ans . Fin de DUA : 01-01-2017
	Communicabilité : AR038 . Librement communicable à partir du : 01-01-2015.
	Taille des documents : 2.8 Mo
	Dates extrêmes : du [sd] au [sd]
	Date création : le 28-08-2015 à 12:44:33
	▶ Autres informations

Figure 35 : Versement accepté visible dans le registre des entrées du SAE

Notre test de versement s'est bien déroulé. Nous avons versé une archive qui contient nos deux fichiers.

IV.3.2 Versement fichier PDF version 1.5 et docx

Dans ce test, nous allons effectuer un versement contenant deux fichiers, un fichier PDF en version 1.5 et un fichier .docx, créé sous Word 2010. Ce versement est effectué depuis l'interface Web d'As@lae.

Une fois le versement créé, le contrôle effectué par le SAE nous informe qu'il y a des erreurs, comme le montre l'image ci-dessous :

Vue détaillée du transfert : #A_CALCULER_LORS_DU_VERROUILLAGE#	
Transfert	Pièces jointes
	Erreur(s) détectée(s)
	Info antivirus
Pièce jointe 'version_1.5.pdf'	
Fichier	version_1.5.pdf [Télécharger]
Taille	1.1 Mo (1 181 093 Octets)
Validation	effectuée par l'outil de validation 'CINES' le 02-09-2015 à 14:10:56
Résultat validation	bien formé : Non, valide : Non, archivable : Non
Info. issues de la validation	format : PDF, version : 1.5
Pièce jointe 'mon_docx.docx'	
Fichier	mon_docx.docx [Télécharger]
Taille	12.3 Ko (12 625 Octets)
Validation	effectuée par l'outil de validation 'CINES' le 02-09-2015 à 14:10:58
Info. issues de la validation	format du document non reconnu, validation non effectuée
Format	le format n'a pas pu être déterminé

Figure 36 : Versement fichier PDF 1.5 et fichier docx non accepté par As@lae

Le fichier PDF 1.5 est reconnu, mais n'est pas archivable du fait qu'il ne fait pas partie des formats autorisés dans l'accord de versement que nous avons créé au préalable.

Le document Word .docx n'est quand à lui pas reconnu. As@lae permet en effet uniquement les formats jusqu'à la version 2003 de Microsoft Word.

Toutefois, en ajoutant un fichier généré sous Word 2003 (.doc), le format semble ne pas être non plus reconnu par As@lae.

IV.3.3 Versement fichier PDF version 1.2

En testant le versement d'un fichier PDF en version 1.2, As@lae autorise l'archivage du document, alors que ce format n'a pas été défini dans l'accord de versement utilisé.

Fichier	version_1.2.pdf [Télécharger]
Taille	323.1 Ko (330 877 Octets)
Validation	effectuée par l'outil de validation 'CINES' le 20-08-2015 à 13:55:14
Résultat validation	bien formé : Oui, valide : Oui, archivable : Oui
Info. issues de la validation	format : PDF, version : 1.2
Jeton d'horodatage	version_1.2.pdf.rep.tsa [Télécharger]

Figure 37 : Versement d'un fichier PDF en version 1.2 sur As@iae

IV.3.4 Autres tests effectués

Nous avons effectué quelques tests fonctionnels supplémentaires via l'interface d'As@iae. Une demande de communication effectuée par un utilisateur fonctionne correctement, après validation par le service de contrôle.

Par contre, le test de demande d'élimination d'une archive dont la DUA était dépassée n'a pas été concluant. Le SAE ne répond pas lorsque le service de contrôle valide cette demande.

IV.4 Test versement automatique depuis la GED Alfresco

Pour tester le versement automatique, nous souhaitons récupérer un fichier présent sur notre GED Alfresco, et l'envoyer via un web service REST d'As@iae vers la maquette du SAE. Il faudra en même temps lire certaines métadonnées du fichier en question afin de générer le bordereau de transfert SEDA.

La meilleure approche serait de développer un bouton action sur la GED qui ferait appel au web service REST en générant en même temps le bordereau SEDA avec les propriétés du document. Cette façon de faire est réalisable, toutefois, appeler un web service externe ou générer un fichier XML nécessite une maîtrise du développement Alfresco qui prendrait du temps à acquérir.

Nous avons décidé par conséquent de découper ce processus en deux étapes. Dans un premier temps, nous avons développé sur Alfresco les métadonnées nécessaires à la génération du bordereau SEDA, création que nous décrirons plus loin.

La deuxième partie sera effectuée à travers un connecteur entre la GED et le SAE, développé en C# avec l'IDE (*Integrated Development Environment*) Visual Studio 2013. C'est ce programme qui effectuera l'appel du web service REST d'As@iae. Avant cet envoi, il sera chargé de récupérer les documents sur la GED, de lire ses métadonnées et de générer le

bordereau de transfert SEDA. Cette passerelle sera lancée manuellement, ou pourra être exécuté en tâche planifiée à une certaine fréquence.

Nous allons maintenant présenter ces deux développements.

IV.4.1 Ajouts de métadonnées sur la GED Alfresco

Tout d'abord, il a fallu créer quelques nouvelles métadonnées qui soient accessibles sur la GED Alfresco (version 5.1). Cela a été fait par le développement d'un aspect personnalisé. Même si cela sort un peu du cadre du mémoire, nous expliquons brièvement ce développement.

Un aspect Alfresco correspond à un ensemble de propriétés que l'on peut appliquer à un document. Ces propriétés sont alors accessibles et définissables sur le document.

Ce développement a été réalisé sous l'IDE Eclipse. Il ne contient que deux fichiers XML (plus le fichier **build.xml**) positionnés dans une arborescence de répertoires correspondant à l'arborescence d'Alfresco :

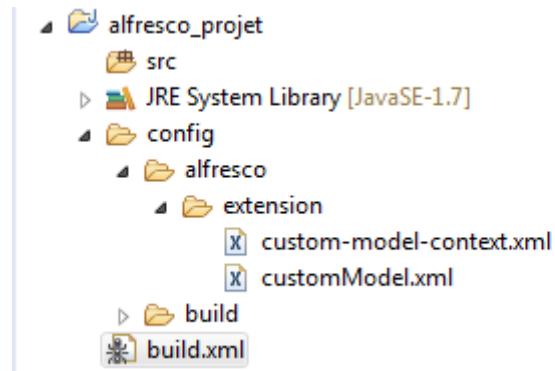


Figure 38 : Arborescence des répertoires correspondant à celle d'Alfresco sous Eclipse

Le fichier '**customModel.xml**' qui contient les propriétés qui seront ajoutées à Alfresco est accessible à l'Annexe 5 : *Fichier pour création d'un aspect Alfresco*.

Sous Eclipse, il faudra ensuite compiler (à partir du fichier **build.xml**) le projet en .jar et le déposer dans un répertoire d'Alfresco. Il faut également configurer quelques fichiers de paramétrage d'Alfresco afin de rendre visible l'aspect (et de le nommer 'Archivable') créé et ses propriétés.

Après avoir redémarré Alfresco, nous pouvons maintenant appliquer notre aspect à un document, memoire.pdf, document en version PDF 1.4.

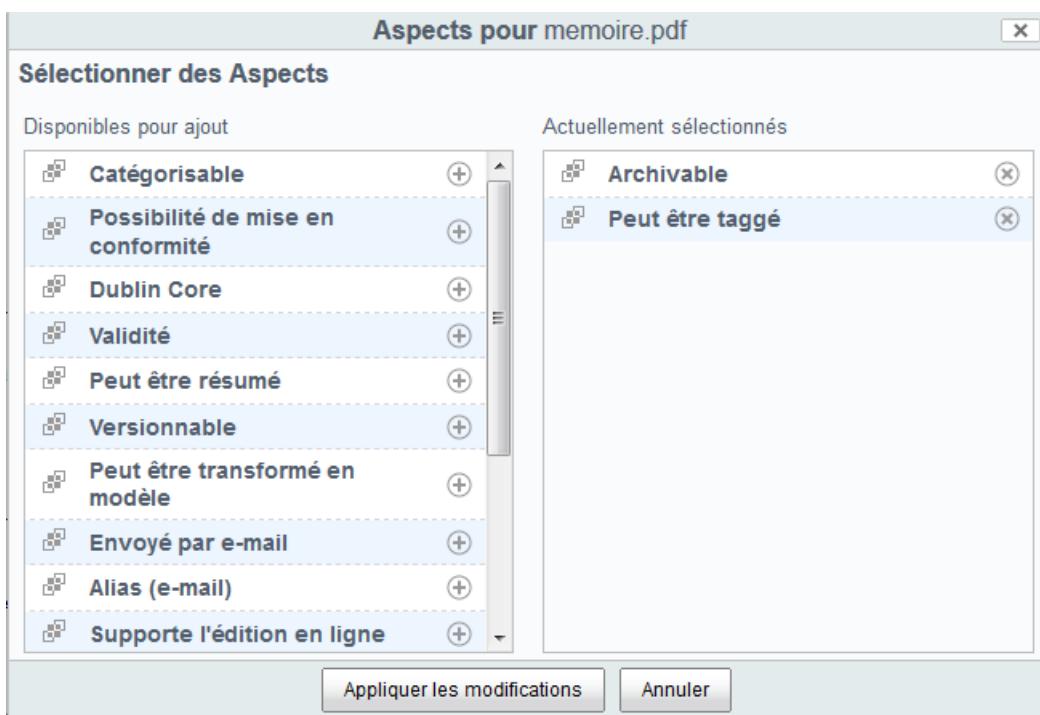


Figure 39 : Ajout de l'aspect personnalisé 'Archivage' sur le document

Nous définissons ensuite les propriétés définies sur notre document.



Editer les propriétés

Modifier les propriétés : memoire.pdf

* Champs requis

Nom: *
memoire.pdf

Titre:
Étude et réalisation d'un prototype d'archivage électronique appliqu

Description:
Archivage électronique dans les collectivités territoriales

Type MIME:
Adobe PDF Document

Auteur:
MONTEL Julien

Tags:

Archiver vers SAE ?

Sort final:

Durée à appliquer (DUA) en années:
2

Règle à appliquer pour communicabilité:
AR039 (25 ans)

Date début de DUA:
1/7/2015

Date début pour communicabilité:
1/7/2015

Objectif:
ceci est l'objectif du document

Commentaire:
document pour tester envoi et récup' métadonnées

Figure 40 : Définition des propriétés qui seront utilisées dans le bordereau SEDA

On retrouve bien les propriétés que nous avons créées dans notre aspect. La propriété « Archiver vers SAE ? » servira à indiquer à la passerelle que le document est à transférer vers As@lae.

IV.4.2 Connecteur pour transférer le document de la GED au SAE

Le deuxième développement effectué est le connecteur qui nous permettra de récupérer les documents à archiver sur la GED et de les transférer vers As@lae.

Tout d'abord la passerelle se connectera à la GED, via le protocole CMIS, et récupérera les documents ayant la métadonnée « À archiver vers SAE » de définie.

Les autres métadonnées du document seront lues et permettront au connecteur de générer le bordereau de transfert au format SEDA.

Toutefois, certaines informations sont écrites en dur dans le bordereau SEDA, comme par exemple l'identifiant du service producteur. As@lae ne propose en effet pas de web service d'interrogation pour récupérer certaines des informations nécessaires pour le bordereau.

Ensuite, du fait qu'As@lae n'autorise que des documents compressés au format .zip pour le transfert d'archive via le web service REST, la passerelle effectuera cette compression si besoin.

Après cette vérification, elle se connectera au web service REST d'As@lae pour transférer le bordereau SEDA et les archives zippées.

Enfin, pour terminer, en cas de succès du versement, elle déplacera le document sur la GED vers un autre répertoire qui aura des droits d'accès restreints. En cas d'erreur, elle générera une alerte.

Le schéma ci-dessous présente succinctement ce cheminement d'étapes.

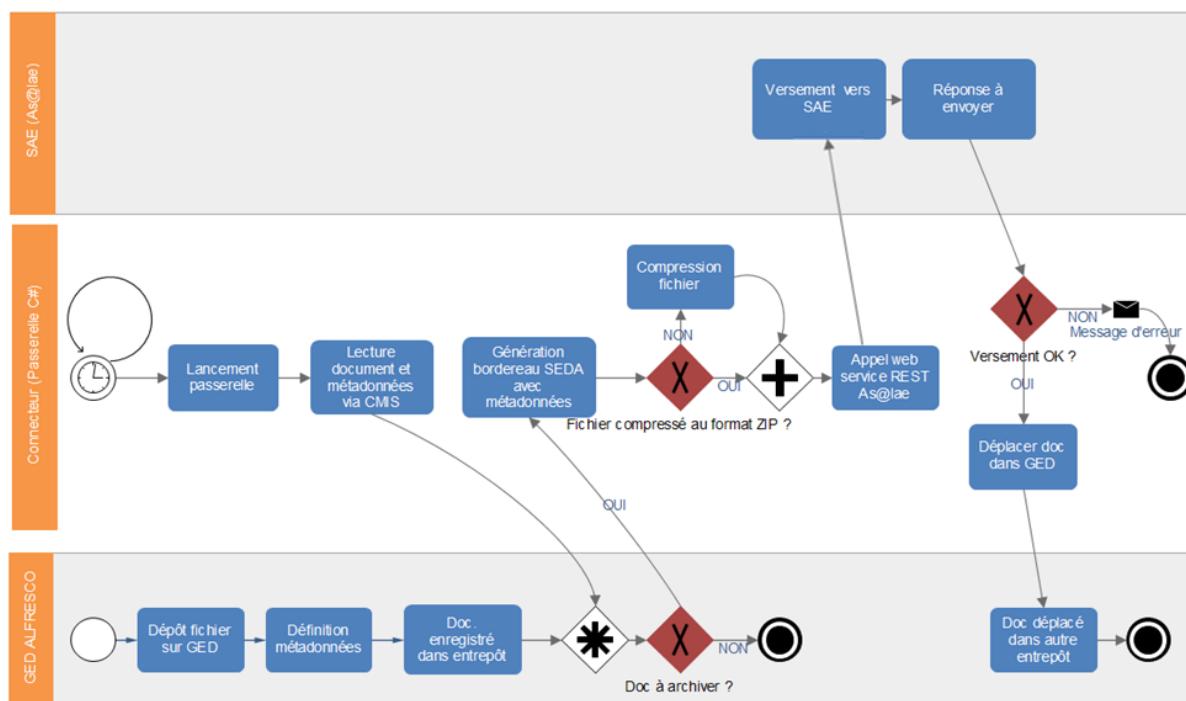


Figure 41 : Diagramme de séquence entre la passerelle, le SAE et la GED

Le code de cette passerelle est présenté en Annexe 6 : *Extrait du code pour le connecteur entre la GED et As@lae*

Le test effectué après lancement de la passerelle est concluant. On retrouve bien notre versement sur As@lae, avec l'indication que le transfert s'est effectué par web service.

Vue détaillée du transfert : MA_23

Transfert Pièces jointes Info antivirus Echanges

Informations internes

Identifiant interne (id)	87
Statut de traitement	En cours de traitement
Conforme	Oui
Décision	En attente
Sens	Entrant
Origine de la création	Transféré par webservice
Date réception	le 02-09-2015 à 21:53:50
Bordereau de transfert	/data/echanges/entree/ArchiveTransfer/87/bordereau.xml
Version du SEDA	0.2
Taille des pièces jointes	3.6 Mo (3 793 699 Octets)

Eléments du bordereau

Service d'archives	Service d'archives
Service versant	Service versant DSI
Service producteur	Service versant DSI
Commentaire	document pour tester envoi et récup' métadonnées

Figure 42 : Détail du transfert par web service REST vers As@lae

IV.5 Évaluation du prototype

Ces tests effectués nous montrent que la maquette est fonctionnelle et répond bien à notre besoin. Nous avons tout de même rencontré quelques dysfonctionnements lors de ces tests.

Concernant les formats de fichiers, les documents Microsoft Word ne sont jamais reconnus, alors qu'As@lae propose ce choix dans la liste des formats autorisés pour un accord de versement.

Pour rappel, As@lae utilise l'outil FormatValidator du CINES, qui propose également un autre outil en ligne, FACILE, permettant de vérifier certains formats de fichiers archivables sur leur plateforme. L'outil FACILE ne permet pas de vérifier les formats de document Microsoft Word. Ne sachant pas si les deux outils sont différents ou pas, il est possible que

les documents Word ne soient pas du tout identifiés par l'outil de validation. Dans ce cas, le SAE ne devrait pas le proposer dans sa liste.

Autre problème concernant les formats, un fichier PDF en version 1.2 est autorisé par le SAE même si ce format n'est pas autorisé dans l'accord de versement. Dans cet exemple, le format est bien reconnu, mais le SAE ne devrait pas l'accepter.

Nous n'avons également pas réussi à valider une demande d'élimination d'un document archivé.

Concernant ces différents problèmes, n'ayant pas de contrat de support Adullact pour la solution As@lae, nous n'avons pas pu pousser plus loin nos investigations.

La passerelle d'interconnexion entre la GED Alfresco et le SAE As@lae est améliorable. Par exemple, comme l'a fait l'équipe de Patrick Pecot du CDG56, il faudrait générer le bordereau SEDA en allant lire le fichier RelaxNG ayant servi à la création du profil d'archivage.

Il faudrait également produire ce bordereau en récupérant toutes les données directement dans As@lae (en plus des métadonnées récupérées sur le document de la GED).

Malheureusement As@lae ne propose pas de web service pour récupérer des informations tels que les identifiants d'acteurs SEDA, les profils d'archives, les accords de versement ou encore les codes utilisés pour la communicabilité de l'archive.

Il est possible d'aller lire directement la base de données As@lae pour récupérer ces informations, mais nous créerions alors un couplage fort entre la passerelle et As@lae. Lors des futures mises à jour du SAE, nous nous exposerions alors à devoir redévelopper la passerelle.

En effet, As@lae n'est actuellement pas réellement basé sur une architecture orientée service (*SOA*¹), mais peut-être est-ce un axe futur pour l'Adullact.

Le tableau ci-dessous synthétise les fonctionnalités d'As@lae, aussi bien celles que nous avons testées que celles qui, mêmes si elles ne sont pas mentionnées, ont été visualisées sur l'interface.

¹ *SOA : Service Oriented Architecture*

Tableau VII : Synthèse des fonctionnalités testées ou pas d'As@lae et résultats des tests

Fonctionnalités	Résultat des tests	Commentaires
Gestion des droits utilisateurs	Oui	
Création de profils d'archives	Oui	
Création d'accords de versement	Oui	
Horodatage des fichiers	Oui	
Calcul empreintes sur fichier	Oui	
Journalisation	Oui	
Gestion des formats de fichiers autorisés	Nuancé	<ul style="list-style-type: none"> - Documents Word non reconnu - Certains formats sont acceptés alors qu'ils ne le devraient pas
Conversion des formats	Test impossible	<ul style="list-style-type: none"> - Les deux outils possibles ne fonctionnent pas : composants mal installés ?
Versement manuel	Oui	
Versement via web service	Oui	
Définition des métadonnées	Oui	
Interrogation du SAE (code, identifiant...)	Nuancé	<ul style="list-style-type: none"> - Uniquement via la base de données Postgresql, et non en web service
Demande de communication	Oui	
Demande de restitution	Non testé	
Demande d'élimination	Non	<ul style="list-style-type: none"> - Fonctionnalité présente mais impossible de la faire fonctionner
Système réparti	Non testé	
Gestion de plusieurs types de stockage	Non testé	
Migration de volume	Non testé - nuancé	<ul style="list-style-type: none"> - Fonctionnalité non visualisée sur l'interface
Recherche par métadonnées	Oui	

Malgré les quelques dysfonctionnements mentionnés, peut-être dus entre autres à de mauvaises manipulation de notre part, la solution As@lae reste un outil de qualité qui offre les différentes fonctionnalités attendues de la part d'un SAE.

Mais comme nous l'avons dit précédemment, ce ne sera peut-être pas la solution choisie lors de la future acquisition d'un SAE patrimonial par Rennes Métropole et la Ville de Rennes.

Conclusion

Pour conclure ce mémoire, l'archivage électronique apparaît comme un domaine en plein essor qui va devenir une nécessité pour les entreprises et les collectivités.

Les collectivités Rennes Métropole et Ville de Rennes n'échappent pas à cette règle et vont devoir se doter d'un SAE patrimoniale dans les prochaines années afin de conserver les documents numériques générés ou reçus dans le cadre de leurs missions.

L'archivage électronique est un domaine vaste et complexe, et ses différentes contraintes, qu'elles soient juridiques, techniques ou normatives rendent le choix d'une solution complexe à effectuer. Toutefois, le domaine normatif commence à reposer sur des bases solides.

La mise en place d'un tel système nécessite également de bien définir les formats de fichiers jugés pérennes, les technologies de supports de stockage résistants.

À travers ce document et le cadre de l'étude menée, nous avons fait le choix de prototyper la solution As@iae de l'Adullact, qui répond relativement bien aux différents critères fixés. Il faut néanmoins garder à l'esprit que du fait du jeune âge de l'archivage électronique, de nouveaux prestataires émergent et proposent de nouvelles solutions de SAE. Par conséquent, le choix d'un SAE doit se faire au démarrage du projet après avoir défini les besoins dans un cahier des charges.

L'État proposera peut-être également une alternative à destination des collectivités territoriales à travers le projet VITAM, projet qu'il faudra par conséquent suivre de près.

Quel que soit le choix qui sera effectué par nos collectivités, les problématiques liées à l'archivage électronique évolueront certainement dans les prochaines années.

L'ère du Big Data et des objets connectés, à travers les concepts des Smart Cities (ou Villes Intelligentes) ou encore des réseaux sociaux d'entreprise (*RSE*), générera en effet de plus en plus de données, qu'il faudra probablement archiver.

Bibliographie

1. Sources et citations

- [1] *D. d. I. I. e. a. Premier Ministre*, «*Quels sont les délais de conservation des documents pour les entreprises ?*», 19 11 2014. [En ligne]. Available: <http://vosdroits.service-public.fr/professionnels-entreprises/F10029.xhtml>. [Accès le 24 11 2014].
- [2] *Direction des Archives de France, (DAF)*, «*Dictionnaire de terminologie archivistique*», 2007. [En ligne]. Available: <http://www.archivesdefrance.culture.gouv.fr/static/3226>. [Accès le 26 11 2014].
- [3] *Locarchives*, «*Qu'est-ce qu'une archive ?*» [En ligne]. Available: <http://www.locarchives.fr/vos-besoins/pourquoi-archiver/quest-ce-quune-archive>. [Accès le 25 11 2014].
- [4] *Direction des Archives de France, (DAF)*, «*Historique des archives*» [En ligne]. Available: <http://www.archivesdefrance.culture.gouv.fr/archives-publiques/historique/>. [Accès le 15 02 2015].
- [5] *F. Papy*, «*Les défis de la gestion documentaire face à la dématérialisation des documents - La théorie des trois âges et le continuum*», chez *Recherches ouvertes sur le numérique*, Hermès - Lavoisier, 2013, pp. 80-83.
- [6] *H. Girard*, «*Projet de loi Crédit : ce que les archivistes veulent obtenir*», *La gazette des communes*, 31 08 2015. [En ligne]. Available: <http://www.lagazettedescommunes.com/385247/projet-de-loi-creation-ce-que-les-archivistes-veulent-obtenir/>. [Accès le 31 08 2015].
- [7] *Secrétariat Général pour la Modernisation de l'Action Publique, (SGMAP)*, «*RGI - interopérabilité*», 12 06 2009. [En ligne]. Available: <http://references.modernisation.gouv.fr/rgi-interoperabilite>. [Accès le 19 03 2015].
- [8] *Agence Nationale de Sécurité des Systèmes d'Information, (ANSSI)*, «*Le Référentiel général de sécurité (RGS)*», 13 06 2014. [En ligne]. Available: <http://www.ssi.gouv.fr/administration/reglementation/administration-electronique/le-referentiel-general-de-securite-rgs/>. [Accès le 19 03 2015].
- [9] *Secrétariat Général pour la Modernisation de l'Action Publique, (SGMAP)*, «*RGAA - Accessibilité*», 2009. [En ligne]. Available:

<https://references.modernisation.gouv.fr/rgaa-accessibilite>. [Accès le 04 2015].

[10] International Organization for Standardization, (ISO), «Normes,» [En ligne]. Available: <http://www.iso.org/iso/fr/home/standards.htm>. [Accès le 03 2015].

[11] F. Banat-Berger et C. (. Huc, «Section 5 : Modèle de référence OAIS et normes associées,» novembre 2011. [En ligne]. Available: http://www.piaf-archives.org/espace-formation/file.php/9/section05_web/co/module_section05.html. [Accès le 04 2015].

[12] Consultative Committee for Space Data Systems, (CCSDS), «Modèle de référence pour un Système d'archivage ouvert (OAIS),» mars 2005. [En ligne]. Available: http://pin.association-aristote.fr/lib/exe/fetch.php/public/documents/norme_oais_version_francaise.pdf. [Accès le 04 2015].

[13] P. Ballet et A. Bensoussan, «Une nouvelle norme pour l'archivage électronique,» 07 05 2009. [En ligne]. Available: <http://www.usinenouvelle.com/article/une-nouvelle-norme-pour-l-archivage-electronique.N66335>.

[14] M.-A. Chabin, «ARCHIVAGE ÉLECTRONIQUE : PROBLÉMATIQUES ET NORMES,» [En ligne]. Available: <http://catalogue-intd.cnam.fr/GEIFFile/10534chabin.pdf?Archive=191310291959&File=10534chabin.pdf>. [Accès le 04 2015].

[15] Direction des Archives de France, (DAF), «ICA-Req,» [En ligne]. Available: <http://www.archivesdefrance.culture.gouv.fr/gerer/archives-electroniques/standard/ica-req/>. [Accès le 04 2015].

[16] Bibliothèque Nationale de France, (BnF), «EAD : Encoded Archival Description,» [En ligne]. Available: http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_ead.html. [Accès le 20 03 2015].

[17] Direction des Archives de France (DAF), «Métadonnées,» [En ligne]. Available: <http://www.archivesdefrance.culture.gouv.fr/gerer/archives-electroniques/standard/metadonnees/>. [Accès le 04 2015].

[18] Library of Congress (Loc), «METS - Metadata Encoding & Transmission Standard,» [En ligne]. Available: <http://www.loc.gov/standards/mets/>. [Accès le 05 2015].

[19] Direction des Archives de France (DAF), «Le standard d'échange de données pour l'archivage (SEDA),» [En ligne]. Available: <http://www.archivesdefrance.culture.gouv.fr/gerer/archives->

electroniques/standard/seda/. [Accès le 05 2015].

[20] Centre National Informatique de l'Enseignement Supérieur (CINES), «Liste des formats archivables,» 05 01 2014. [En ligne]. Available: <https://www.cines.fr/archivage/des-expertises/expertise-formats/liste-des-formats-archivables/>. [Accès le 05 2015].

[21] Direction Interministérielle des Systèmes d'Information et de Communication (DISIC), «Archivage électronique - Un nouveau domaine d'expertise au service de la gouvernance des systèmes d'information,» 11 09 2012. [En ligne]. Available: http://references.modernisation.gouv.fr/sites/default/files/DISIC_AE%20Guide%20bonnes%20pratiques.pdf. [Accès le 05 2015].

[22] Gurumed, «Les cristaux de mémoires pourraient devenir une réalité,» 15 08 2011. [En ligne]. Available: <http://www.gurumed.org/2011/08/15/les-cristaux-de-mmoire-pourraient-devenir-une-ralit/>. [Accès le 06 2015].

[23] A. Bastide, «Le quartz, support de stockage pour des milliers d'années,» 24 10 2012. [En ligne]. Available: <http://www.indexel.net/actualites/le-quartz-support-de-stockage-pour-des-milliers-d-annees-3674.html>. [Accès le 06 2015].

[24] L. Cailloux, «Des molécules pour stocker l'information,» 20 04 2015. [En ligne]. Available: <https://lejournal.cnrs.fr/articles/des-molecules-pour-stocker-linformation>.

[25] J. Cheminat, «L'ADN, une solution pour le stockage de demain,» 15 01 2013. [En ligne]. Available: <http://www.lemondeinformatique.fr/actualites/lire-l-adn-une-solution-pour-le-stockage-de-demain-52238.html>.

[26] Groupement d'Intérêt Scientifique sur les Supports Pérennes d'Archivage des DONnées Numériques (GIS-DON devenu GIS-SPADON), «Recommandations du GIS-DON concernant la préservation à long terme des données numériques,» [En ligne]. Available: <https://www.lne.fr/publications/recherche/recommandations-gis-don-conservation-donnees.pdf>. [Accès le 05 2015].

[27] Agence Nationale de Sécurité des Systèmes d'Information, (ANSSI), «Cahier des charges pour un système d'archivage électronique,» 16 05 2006. [En ligne]. Available: <http://www.ssi.gouv.fr/uploads/2015/02/ArchivageSecurise-CahierDesCharges-2006-05-16.pdf>. [Accès le 06 2015].

[28] Secrétariat Général pour la Modernisation de l'Action Publique, (SGMAP), «Vitam : Vers un socle d'archivage électronique commun à toute l'administration,» [En ligne]. Available: <http://www.modernisation.gouv.fr/ladministration-change-avec-le-numerique/par-son-systeme-dinformation/vitam-vers-un-socle-d-archivage->

electronique-commun-toute-l-administration. [Accès le 25 08 2015].

- [29] P. (. o. C. Caplan, «Comprendre PREMIS,» 01 02 2009. [En ligne]. Available: http://www.loc.gov/standards/premis/Understanding-PREMIS_french.pdf. [Accès le 04 2015].
- [30] Conseil International des Archives, (CIA), «ISAD(G): Norme générale et internationale de description archivistique - Deuxième édition,» [En ligne]. Available: <http://www.ica.org/10225/normes/isadg-norme-gnrale-et-internationale-de-description-archivistique-deuxime-dition.html>. [Accès le 20 03 2015].
- [31] Portail International Archivistique Francophone, (PIAF), «Publication électronique des instruments de recherche - ISAD(G) et EAD,» [En ligne]. Available: http://www.piaf-archives.org/espace-formation/file.php/8/06section8_web_web/co/06_section8_27.html. [Accès le 20 03 2015].
- [32] Direction des Archives de France, (DAF), «Normes nationales et internationales,» [En ligne]. Available: <http://www.archivesdefrance.culture.gouv.fr/gerer/classement/normes-outils/nationales-internationales/>. [Accès le 20 03 2015].
- [33] Rennes Métropole, «4121 habitants de plus dans l'agglomération,» Rennes Métropole, 06 01 2014. [En ligne]. Available: <http://metropole.rennes.fr/actualites/institutions-citoyennete/institution/4121-habitants-de-plus-dans-l-agglomeration/>.
- [34] D. (. Boucon, «PAIMAS et PAIS,» 07 10 2010. [En ligne]. Available: <http://pin.association-aristote.fr/lib/exe/fetch.php/public/presentations/2010/pin20101007-pres04-paimaspais.pdf>. [Accès le 04 2015].
- [35] Futura Sciences, «Le disque d'Arnano archive ses données pour des siècles,» 27 06 2012. [En ligne]. Available: <http://www.futura-sciences.com/magazines/high-tech/infos/actu/d/informatique-disque-arnano-archive-donnees-siecles-39638/>. [Accès le 05 2015].

2. Lectures pour aller plus loin

- <http://www.modernisation.gouv.fr/ladministration-change-avec-le-numerique/par-son-systeme-dinformation/vitam-vers-un-socle-d-archivage-electronique-commun-toute-l-administration>
- <http://www.fntc.org/publications/les-guides/les-guides-de-la-confiance/>
- <http://references.modernisation.gouv.fr/archivage-numerique>
- <http://siaf.hypotheses.org/pour-en-savoir-plus>
- <http://siafdroit.hypotheses.org/>
- <http://www.archivesdefrance.culture.gouv.fr/gerer/>
- https://www.zotero.org/groups/archives_de_france - base_juridique_pour_la_collecte_des_archives
- <http://www.archivesdefrance.culture.gouv.fr/seda/>
- <http://www.archivistes.org/>
- <http://www.piaf-archives.org/>
- <https://normesarchivage.wordpress.com/>
- <http://www.archive17.fr/index.php/Normes-et-guides/normes-et-guides-records-management-preservation.html>
- <http://www.guideinformatique.com/dossiers-actualites-informatiques-gestion-commerciale-4/archivage-legal-233.html>
- <https://archivesonline.wordpress.com/2010/12/30/le-module-2-dica-req-exigences-fonctionnelles-pour-les-sae/>

Liste des figures

Figure 1 : Schéma illustrant les différentes briques du projet dans sa globalité	14
Figure 2 : Diagramme de GANTT - Planning prévisionnel du projet de mémoire	16
Figure 3 : Cartographie des communes de Rennes Métropole	22
Figure 4 : Organigramme de la Direction des Systèmes d'Information	27
Figure 5 : Rennes, ville capitale de Bretagne et siège du Parlement 1616.....	34
Figure 6 : Théorie des 3 âges - cycle de vie du document	37
Figure 7 : Image humoristique sur la numérisation des archives	39
Figure 8 : Modèle du paquet d'information archivé (AIP)	45
Figure 9 : Modèle fonctionnel OAIS	46
Figure 10 : Architecture découpée en services d'un SAE conforme MoReq2010	49
Figure 11 : Exemple de texte visualisé dans Word 2010.....	56
Figure 12 : Exemple d'un document Word 2010 visualisé avec Notepad++	57
Figure 13: Disque nanoforme de la société Arnano.....	61
Figure 14 : Modèle fonctionnel OAIS	67
Figure 15 : Diagramme opérationnel de contexte du système de SAE patrimonial.....	68
Figure 16 : Diagramme des tâches des différents acteurs du SAE.....	69
Figure 17: Architecture cible du futur système.....	76
Figure 18 : Génération d'un certificat	84
Figure 19 : Paramétrage du fichier jhove.conf.....	85
Figure 20 : Architecture d'As@lae.....	87
Figure 21 : Écran de connexion à l'application As@lae	88
Figure 22 : Page d'accueil d'As@lae	89
Figure 23 : Création d'un acteur SEDA sous As@lae	90
Figure 24 : Génération d'un fichier RelaxNG avec l'outil Agape	91
Figure 25 : Création d'un profil d'archivage sous As@lae	92
Figure 26 : Création d'un accord de versement sous As@lae	93
Figure 27 : Définition des formats de fichiers autorisés pour un accord de versement sous As@lae.....	93
Figure 28 : Page d'accueil d'As@lae avec des droits restreints	94
Figure 29 : Versement manuel d'archive depuis As@lae : entête du transfert	95
Figure 30 : Versement manuel, intitulé de l'archive et choix du profil d'archivage et accord de versement.....	95
Figure 31 : Ajout du document PDF en tant que contenu de données	96
Figure 32 : Ajout du fichier de métadonnées METS.....	96
Figure 33 : Visualisation des pièces jointes ajoutées lors de la création du versement.....	97
Figure 34 : Validation du versement par le service des Archives	98
Figure 35 : Versement accepté visible dans le registre des entrées du SAE.....	98
Figure 36 : Versement fichier PDF 1.5 et fichier docx non accepté par As@lae	99
Figure 37 : Versement d'un fichier PDF en version 1.2 sur As@lae.....	100

Figure 38 : Arborescence des répertoires correspondant à celle d'Alfresco sous Eclipse.....	101
Figure 39 : Ajout de l'aspect personnalisé 'Archivage' sur le document	102
Figure 40 : Définition des propriétés qui seront utilisées dans le bordereau SEDA	103
Figure 41 : Diagramme de séquence entre la passerelle, le SAE et la GED	104
Figure 42 : Détail du transfert par web service REST vers As@lae	105

Liste des tableaux

Tableau I : Matrice de cotation des risques	18
Tableau II : Tableau récapitulatif des normes en matière d'archivage électronique	54
Tableau III : Estimation de la durée de vie des supports d'archivage électronique	60
Tableau IV : Fonctionnalités comparées SAE / GED / CFN	66
Tableau V : Liste des exigences du SAE	73
Tableau VI : Comparatif des fonctionnalités des solutions SAE.....	79
Tableau VII : Synthèse des fonctionnalités testées ou pas d'As@lae et résultats des tests .	107

Liste des annexes

Annexe 1 : Exemple PREMIS.....	118
Annexe 2 : Exemple METS	122
Annexe 3 : Fichier RelaxNG pour profil d'archives.....	124
Annexe 4 : Bordereau de transfert SEDA	125
Annexe 5 : Fichier pour création d'un aspect Alfresco	127
Annexe 6 : Extrait du code pour le connecteur entre la GED et As@iae.....	130

Annexe 1 : Exemple PREMIS

L'exemple ci-dessous a été récupéré en intégralité sur la traduction française de la documentation de PREMIS [18]¹.

« L'exemple ci-dessous est la simplification d'un exemple conçu par la Bibliothèque du Congrès pour un tutoriel PREMIS. Il montre les unités sémantiques PREMIS et les valeurs utilisées pour décrire une image TIFF.

La LoC utilise des identifiants handle (identifiants créés et gérés par une application appelée le système Handle) pour les fichiers stockés dans l'entrepôt (cf. 1.1 objectIdentifier et 1.10.3 relatedObjectIdentification). L'objet décrit est reconnu dans l'entrepôt par le biais de son identifiant unique (1.1 objectIdentifier) mais son nom d'origine avant le versement était « 001h.tif » (1.6 originalName). Nous savons qu'il s'agit d'un fichier grâce à la catégorie de l'objet (1.2 objectCategory).

Le fichier possède une empreinte MD5 calculée par le système de gestion de contenus de la LoC (1.5.2 fixity). Son format est décrit de deux manières différentes, par un type MIME et un numéro de version (1.5.4.1 formatDesignation) et en pointant vers le répertoire de formats PRONOM qui détient une copie de la spécification TIFF.

La LoC considère qu'il y a eu deux applications d'origine pour ce fichier. Un élément non déterminé (probablement un fichier RAW sans nom) a été créé sur un scanner par l'application ScandAll et a été transformé en TIFF en utilisant Adobe Photoshop (1.5.5 creatingApplication). Il n'y a pas de cryptage ou d'autre type de restriction que la LoC ait jugé pertinent de consigner explicitement (1.5.6 inhibitors).

Le fichier est stocké sur disque à un emplacement identifié par son chemin dans l'arborescence (1.7.1 contentLocation) ; le type de support correspondant est décrit très précisément, sans doute afin que les gestionnaires de l'entrepôt puissent déterminer le moment où les données nécessiteront une migration vers un type de support plus récent (1.7.2 storageMedium).

La LoC a pris la peine de consigner un environnement recommandé pour visualiser le fichier : utiliser Adobe Acrobat version 5.0 sur un processeur Intel x86 fonctionnant sous Windows XP.

La vitesse du processeur doit être au moins de 60 MHz et il doit y avoir au moins 32 Mo de mémoire, et de préférence 64 Mo (1.8 environment). Le fichier TIFF fait partie d'une représentation dont l'identifiant interne est R200802948.

Cette information est donnée par le troisième bloc d'informations sur les relations (1.10 relationship). Il y a deux autres fichiers dans cette représentation, identifiés par des identifiants Handle. Nous savons qu'il doit s'agir de fichiers car le type de relation consigné est « frère ». Il nous faut consulter la description des objets loc.music/gottlieb.09602 et loc.music/gottlieb.mets pour plus d'informations sur ces derniers, par exemple pour savoir de quel type de fichiers il s'agit.

¹ Library of Congress – Caplan P. : Comprendre PREMIS

Nous n'avons pas d'information descriptive sur l'entité intellectuelle correspondant à cette représentation. Toutefois, nous avons un lien vers son identifiant, un LCCN (1.12 linkingIntellectualEntityIdentifier). En toute logique, si l'on recherchait cet LCCN, nous trouverions une description bibliographique de l'entité intellectuelle en question (pas dans cet exemple toutefois car le LCCN a été forgé). »

Unité sémantique	Valeur
1.1 objectIdentifier (identifiant de l'objet)	
1.1.1 objectIdentifierType (type d'identifiant de l'objet)	hd1
1.1.2 objectIdentifierValue (valeur de l'identifiant de l'objet)	loc.music/gottlieb.09601
1.2 objectCategory (catégorie de l'objet)	file
1.3 preservationLevel (niveau de préservation)	
1.3.1 preservationLevelValue (valeur du niveau de préservation)	full
1.5 objectCharacteristics (caractéristiques de l'objet)	
1.5.1 compositionLevel (niveau de composition)	0
1.5.2 fixity (intégrité)	
1.5.2.1 messageDigestAlgorithm (algorithme de l'empreinte numérique)	MD5
1.5.2.2 messageDigest (empreinte numérique)	36b03197ad066cd719906c55eb68ab8d
1.5.2.3 messageDigestOriginator(origine de l'empreinte numérique)	localDCMS
1.5.3 size (taille)	20800896
1.5.4 format	
1.5.4.1 formatDesignation (dénomination du format)	
1.5.4.1.1 formatName(nom du format)	image/tiff
1.5.4.1.2 formatVersion(verison du format)	6.0
1.5.4.2 formatRegistry (répertoire de formats)	
1.5.4.2.1 formatRegistryName(nom répertoire de formats)	PRONOM
1.5.4.2.2 formatRegistryKey (identifiant du répertoire de formats)	fmt/10
1.5.4.2.3 formatRegistryRole (fonction du répertoire de formats)	specification
1.5.5 creatingApplication (application d'origine)	
1.5.5.1 creatingApplicationName (nom de l'application d'origine)	ScandAll21
1.5.5.2 creatingApplication Version (version de l'application d'origine)	4.1.4

1.5.5.3 dateCreatedByApplication (date de création par l'application d'origine)	1998-10-30
1.5.5 creatingApplication (application d'origine)	
1.5.5.1 creatingApplicationName (nom de l'application d'origine)	Adobe Photoshop
1.5.5.2 creatingApplication Version (version de l'application d'origine)	CS2
1.5.5.3 dateCreatedByApplication (date de création par l'application d'origine)	1998-10-30T08:29:02
1.5.6 inhibitors (restrictions)	
1.5.6.1 inhibitorType (type de restriction)	none
1.6 originalName (nom d'origine)	001h.tif
1.7 storage (stockage)	
1.7.1 contentLocation (emplacement du contenu)	
1.7.1.1 contentLocationType (type d'emplacement d'origine)	filepath
1.7.1.2 contentLocationValue (valeur de l'emplacement d'origine)	amserver/
1.8 environment	
1.8.1 environmentCharacteristic (caractéristique de l'environnement)	recommended
1.8.2 environmentPurpose (fonction de l'environnement)	render
1.8.2 environmentPurpose (fonction de l'environnement)	edit
1.8.5 software (logiciel)	
1.8.5.1 swName (nom du logiciel)	Adobe Acrobat
1.8.5.2 swVersion (version du logiciel)	5.0
1.8.5.3 swType (type de logiciel)	renderer
1.8.5 software (logiciel)	
1.8.5.1 swName (nom du logiciel)	Windows
1.8.5.2 swVersion (version du logiciel)	XP
1.8.5.3 swType (type de logiciel)	Operating System
1.8.6 hardware (matériel)	
1.8.6.1 hwName (nom du matériel)	Intel x86
1.8.6.2 hwType (type de matériel)	processor
1.8.6.3 hwOtherInformation (autres informations sur le matériel)	60 Mhz minimum
1.8.6 hardware (matériel)	
1.8.6.1 hwName (nom du matériel)	64 MB RAM
1.8.6.2 hwType (type de matériel)	memory
1.8.6.3 hwOtherInformation (autres informations sur le matériel)	32 MB minimum
1.10 relationship (relation)	
1.10.1 relationshipType (type de relation)	structural
1.10.2 relationshipSubType (sous-type de relation)	has sibling

1.10.3 relatedObjectIdentification (identification de l'objet en relation)	
1.10.3.1 relatedObjectIdentifierType (type de l'identifiant de l'objet en relation)	hd1
1.10.3.2 relatedObjectIdentifierValue (valeur de l'identifiant de l'objet en relation)	loc.music/gottlieb.09602
1.10.3.3 relatedObjectSequence (ordre de l'objet lié)	0
1.10 relationship (relation)	
1.10.1 relationshipType (type de relation)	structural
1.10.2 relationshipSubType (sous-type de relation)	has sibling
1.10.3 relatedObjectIdentification (identification de l'objet en relation)	
1.10.3.1 relatedObjectIdentifierType (type de l'identifiant de l'objet en relation)	hd1
1.10.3.2 relatedObjectIdentifierValue (valeur de l'identifiant de l'objet en relation)	loc.music/gottlieb.mets
1.10.3.3 relatedObjectSequence (ordre de l'objet lié)	0
1.10 relationship (relation)	
1.10.1 relationshipType (type de relation)	structural
1.10.2 relationshipSubType (sous-type de relation)	is included in
1.10.3 relatedObjectIdentification (identification de l'objet en relation)	
1.10.3.1 relatedObjectIdentifierType (type de l'identifiant de l'objet en relation)	LocalRepository
1.10.3.2 relatedObjectIdentifierValue (valeur de l'identifiant de l'objet en relation)	R200802948
1.10.3.3 relatedObjectSequence (ordre de l'objet lié)	0
1.12 linkingIntellectualEntityIdentifier (identifiant de l'entité intellectuelle liée)	
1.12.1 linkingIntellectualEntityIdentifierType (type de l'identifiant de l'entité intellectuelle liée)	LCCN
1.12.2 linkingIntellectualEntityIdentifierValue (valeur de l'identifiant de l'entité intellectuelle liée)	2007-86121

Annexe 2 : Exemple METS

Ci-dessous un exemple de fichier METS pour ce document créé sous Word 2010, une fois transformé en PDF-A. Des métadonnées Dublin Core et PREMIS y sont ajoutées.

```

<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:premis="info:lc/xmlns/premis-v2"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/METS/
    http://www.loc.gov/standards/mets/mets.xsd
    http://www.w3.org/1999/xlink
    http://www.loc.gov/standards/mets/xlink.xsd
    http://purl.org/dc/elements/1.1/
    http://dublincore.org/schemas/xmls/qdc/2003/04/02/dc.xsd
    http://www.loc.gov/standards/premis/v1/Object-v1-1.xsd">
  <!-- Renseignement sur le fichier METS -->
  <mets:metsHdr CREATEDATE="2015-04-15T19:00:00" RECORDSTATUS="Complete">
    <mets:agent ROLE="CREATOR" TYPE="INDIVIDUAL">
      <mets:name>Julien MONTEL</mets:name>
    </mets:agent>
  </mets:metsHdr>
  <!-- Metadonnées descriptives internes, on pourrait également mettre des pointeurs externes -->
  <mets:dmdSec ID="memoire_cnam_sae_jm">
    <!-- type Dublin Core (DC) -->
    <mets:mdWrap MDTYPE="DC">
      <mets:xmlData>
        <dc:creator>Julien MONTEL</dc:creator>
        <dc:title>Étude et réalisation d'un prototype d'archivage électronique appliquée aux collectivités territoriales</dc:title>
        <dc:format>application/pdf</dc:format>
        <dc:date>2015-05-12</dc:date>
        <dc:type>Mémoire électronique d'étude d'ingénieur</dc:type>
        <dc:coverage>France</dc:coverage>
        <dc:coverage>Rennes Métropole</dc:coverage>
        <dc:coverage>Conservatoire National des Arts et Métiers</dc:coverage>
        <dc:coverage>Archivage Électronique</dc:coverage>
        <dc:subject>Archivage Électronique</dc:subject>
        <dc:subject>Collectivités territoriales</dc:subject>
        <dc:subject>SAE</dc:subject>
        <dc:subject>Mémoire d'ingénieur</dc:subject>
        <dc:language>fr</dc:language>
        <dc:identifier>mémoire_v17.pdf</dc:identifier>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
  <!-- Métdonnées Administratives -->
  <mets:amdSec>
    <mets:techMD ID="premis_pdf">
      <mets:mdWrap MDTYPE="PREMIS">
        <mets:xmlData>
          <premis:object>
            <premis:objectIdentifier>
              <premis:objectIdentifierType>Mémoire JM PID

```

```

</premis:objectIdentifierType>
<premis:objectIdentifierValue>uuid:0b83a11e-cdee-47e7-8968-
f99d23a2eba4
    </premis:objectIdentifierValue>
</premis:objectIdentifier>
<objectCharacteristics>
    <fixity>
        <messageDigestAlgorithm>SHA256</messageDigestAlgorithm>
<messageDigest>b6411d3ddb804077d721200452f6312e905fadb1be8973fc85b3f207fac7
2b8d</messageDigest>
        <messageDigestOriginator>Rennes
Métropole</messageDigestOriginator>
    </fixity>
    <size>184302</size>
    <format>
        <formatDesignation>
            <formatName>application/pdf</formatName>
            <formatVersion>PDF-A</formatVersion>
        </formatDesignation>
    </format>
</objectCharacteristics>
<environment>
    <software>
        <swName>Microsoft Word</swName>
        <swVersion>2010</swVersion>
    </software>
    <software>
        <swName>Microsoft Windows</swName>
        <swVersion>7</swVersion>
        <swType>Operating System</swType>
    </software>
</environment>
</premis:object>
</mets:xmlData>
</mets:mdWrap>
</mets:techMD>
</mets:amdSec>
<!-- Description de l'ensemble des fichiers utilisés -->
<mets:fileSec>
    <!-- Fichiers rassemblés par groupes de fichiers -->
    <mets:fileGrp ADMID="premis_pdf">
        <mets:file ID="memoire_v17.pdf" MIMETYPE="application/pdf">
            <mets:Flocat LOCTYPE="URL" xlink:href="memoire_v17.pdf"/>
        </mets:file>
    </mets:fileGrp>
</mets:fileSec>
<!-- Carte de structure -->
<mets:structMap>
    <mets:div>
        </mets:div>
    </mets:structMap>
</ mets:mets>

```

Annexe 3 : Fichier RelaxNG pour profil d'archives

Ci-dessous le début du fichier RelaxNG qui servira pour créer le profil d'archives. Ce fichier a été créé par l'application Agape, et définit le contenu du bordereau de transfert au format SEDA.

```

<?xml version="1.0" encoding="UTF-8"?>
<rng:grammar xmlns:rng="http://relaxng.org/ns/structure/1.0"
  xmlns:udt="urn:un:unece:uncefact:data:standard:UnqualifiedDataType:6"
  xmlns:a="http://relaxng.org/ns/compatibility/annotations/1.0"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:qdt="fr:gouv:ae:archive:draft:standard_echange_v0.2:QualifiedDataType
:1"
  xmlns:ccts="urn:un:unece:uncefact:documentation:standard:CoreComponentsTech
nicalSpecification:2"
  xmlns="fr:gouv:ae:archive:draft:standard_echange_v0.2"
  ns="fr:gouv:ae:archive:draft:standard_echange_v0.2"
  datatypeLibrary="http://www.w3.org/2001/XMLSchema-datatypes">
  <rng:start combine="choice">
    <rng:ref name="ArchiveTransfer"/>
  </rng:start>
  <rng:define name="ArchiveTransfer">
    <rng:element name="ArchiveTransfer">
      <rng:ref name="ArchiveTransfer_N65537"/>
    </rng:element>
  </rng:define>
  <rng:define name="ArchiveTransfer_N65537">
    <rng:optional>
      <rng:element name="Comment">
        <rng:ref name="Comment_N65541"/>
      </rng:element>
    </rng:optional>
    <rng:element name="Date">
      <rng:data type="string"/>
    </rng:element>
    <rng:element name="TransferIdentifier">
      <rng:ref name="TransferIdentifier_N65558"/>
    </rng:element>
    <rng:element name="TransferringAgency">
      <rng:ref name="TransferringAgency_N65588"/>
    </rng:element>
    <rng:element name="ArchivalAgency">
      <rng:ref name="ArchivalAgency_N65632"/>
    </rng:element>
    <rng:zeroOrMore>
      <rng:element name="Integrity">
        <rng:ref name="anyElement"/>
      </rng:element>
    </rng:zeroOrMore>
    [...]
  
```

Annexe 4 : Bordereau de transfert SEDA

Ci-dessous le bordereau de transfert au format SEDA généré par As@lae lors du premier test de versement manuel d'archives.

```

<?xml version="1.0" encoding="UTF-8"?>
<ArchiveTransfer xmlns="fr:gouv:ae:archive:draft:standard_echange_v0.2">
  <Comment>premier transfert de la DSI vers SAE</Comment>
  <Date>2015-08-20T00:00:00+02:00</Date>
  <TransferIdentifier>AT_32</TransferIdentifier>
  <TransferringAgency>
    <Identification>FRAD001</Identification>
  </TransferringAgency>
  <ArchivalAgency>
    <Identification>FRAD000</Identification>
    <Name>Service d'archives départemental</Name>
  </ArchivalAgency>
  <Integrity>
    <Contains
      algorithme="http://www.w3.org/2001/04/xmlenc#sha256">3bf01dc260dceb08dc67a5
      7cdc5918ec151bf554306a9b312c93d9e2bc1763a1</Contains>
      <UnitIdentifier>memoire_v17.pdf</UnitIdentifier>
    </Integrity>
    <Integrity>
      <Contains
        algorithme="http://www.w3.org/2001/04/xmlenc#sha256">662ae2c9017d8927930aa3
        573a15ce2affa1b5cccd52c6a473e83509a12a6fabb</Contains>
        <UnitIdentifier>mets_mem_pdf.xml</UnitIdentifier>
      </Integrity>
      <Contains
        <ArchivalAgreement>ACC_VERS_DSI01</ArchivalAgreement>
        <ArchivalProfile>ID_PROFIL_02</ArchivalProfile>
        <DescriptionLanguage listVersionID="edition
        2009">fr</DescriptionLanguage>
        <DescriptionLevel listVersionID="edition 2009">class</DescriptionLevel>
        <Name>Version actuelle du mémoire JM</Name>
        <ContentDescription>
          <Description>Ceci est la version actuelle du mémoire de
          JM</Description>
          <Language listVersionID="edition 2009">fr</Language>
        </ContentDescription>
        <Appraisal>
          <Code listVersionID="edition 2009">conserver</Code>
          <Duration>P2Y</Duration>
          <StartDate>2015-01-01</StartDate>
        </Appraisal>
        <AccessRestriction>
          <Code listVersionID="edition 2009">AR038</Code>
          <StartDate>2015-01-01</StartDate>
        </AccessRestriction>
        <Document>
          <Attachment filename="memoire_v17.pdf"></Attachment>
        </Document>
      </Contains>
    </Integrity>
  </ArchivalAgency>
</ArchiveTransfer>
  
```

```
<Description>Document du mémoire JM</Description>
<Purpose>Version 17 du mémoire de JM sur l'archivage
électronique</Purpose>
<Type listVersionID="edition 2009">CDO</Type>
</Document>
<Document>
<Attachment filename="mets_mem_pdf.xml"></Attachment>
<Description>Métadonnées METS associées au mémoire version 17 de
JM</Description>
<Purpose>Métadonnées descriptives, de provenance</Purpose>
<Type listVersionID="edition 2009">PDI</Type>
</Document>
</Contains>
</ArchiveTransfer>
```

Annexe 5 : Fichier pour création d'un aspect Alfresco

Ci-dessous le fichier XML créé pour générer un nouvel Aspect Alfresco contenant des propriétés spécifiques à l'archivage.

```

<?xml version="1.0" encoding="UTF-8"?>
<model name="my:custommodel"
       xmlns="http://www.alfresco.org/model/dictionary/1.0">
  <description>Création d'un aspect Alfresco personnalisé pour
l'archivage</description>
  <author>Julien Montel</author>
  <version>1.0</version>
  <!-- Imports are required to allow references to definitions in other
models -->
  <imports>
    <!-- Import Alfresco Dictionary Definitions -->
    <import uri="http://www.alfresco.org/model/dictionary/1.0"
            prefix="d" />
    <!-- Import Alfresco Content Domain Model Definitions -->
    <import uri="http://www.alfresco.org/model/content/1.0"
            prefix="cm" />
  </imports>
  <!-- Introduction of new namespaces defined by this model -->
  <namespaces>
    <namespace uri="http://www.mycompany.com/model/content/1.0"
               prefix="my" />
  </namespaces>
  <constraints>
    <constraint name="my:listSortFinal" type="LIST">
      <parameter name="allowedValues">
        <list>
          <value>Conserver</value>
          <value>Détruire</value>
        </list>
      </parameter>
    </constraint>
    <constraint name="my:listDureeDua" type="LIST">
      <parameter name="allowedValues">
        <list>
          <value>0</value>
          <value>1</value>
          <value>2</value>
          <!-- others values [...] -->
        </list>
      </parameter>
    </constraint>
    <constraint name="my:listRegleComm" type="LIST">
      <parameter name="allowedValues">
        <list>
          <value>AR038 (0 an)</value>
          <value>AR039 (25 ans)</value>
          <!-- others values [...] -->
        </list>
      </parameter>
    </constraint>
  </constraints>
</model>
  
```

```

        </list>
      </parameter>
    </constraint>
  </constraints>
<aspects>
  <aspect name="my:transfertToSAE">
    <title>Propriétés pour archivage vers SAE</title>
    <properties>
      <property name="my:isToArchive">
        <title>Archiver vers SAE ?</title>
        <description>Cocher pour archiver vers le
          SAE</description>
        <type>d:boolean</type>
        <default>false</default>
      </property>
      <property name="my:sortFinal">
        <title>Sort final</title>
        <description>Sort final du document (Conserver,
          Détruire)</description>
        <type>d:text</type>
        <multiple>false</multiple>
        <constraints>
          <constraint ref="my:listSortFinal" />
        </constraints>
      </property>
      <property name="my:dureeSortFinal">
        <title>Durée à appliquer (DUA) en années</title>
        <description>Durée en années avant sort
          final</description>
        <type>d:text</type>
        <multiple>false</multiple>
        <constraints>
          <constraint ref="my:listDureeDua" />
        </constraints>
      </property>
      <property name="my:duaDateDeb">
        <title>Date début de DUA</title>
        <description>Date de début du calcul pour sort
          final (DUA)</description>
        <type>d:date</type>
      </property>
      <property name="my:regleComm">
        <title>Règle à appliquer pour
          communicabilité</title>
        <description>Durée avant
          communicabilité</description>
        <type>d:text</type>
        <multiple>false</multiple>
        <constraints>
          <constraint ref="my:listRegleComm" />
        </constraints>
      </property>
      <property name="my:comDateDeb">

```

```
<title>Date début pour communicabilité</title>
<description>Date de départ du calcul pour la
communicabilité</description>
<type>d:date</type>
</property>
<property name="my:comment">
<title>Commentaire</title>
<description>Commentaire pour
archivage</description>
<type>d:text</type>
</property>
<property name="my:object">
<title>Objectif</title>
<description>Objectif du document
archivé</description>
<type>d:text</type>
</property>my:regleComm
</properties>
</aspect>
</aspects>
</model>
```

Annexe 6 : Extrait du code pour le connecteur entre la GED et As@lae

Nous présentons ci-dessous des extraits du code C# pour le développement effectué du connecteur entre la GED Alfresco et le SAE As@lae. Étant donné la longueur du code, nous n'avons en effet pas mis la totalité du code. Nous ne mettons par exemple pas la déclaration de toutes les variables, la gestion des erreurs ou encore toute la génération du bordereau SEDA.

Je tiens également à remercier Patrick Percot du CDG56¹ (et son équipe) qui m'a fourni le code source de leur programme pour faire appel au web service REST d'As@lae. Le code source complet de leur solution est accessible sur les forges :

<https://github.com/PatPercot/Seda-Generator> et <https://sedaautogenerator.codeplex.com/>

```
[...]Extrait de la méthode de connexion à Alfresco et recuperation des fichiers à archiver
Dictionary<string, string> parameters = new Dictionary<string, string>();
parameters[DotCMIS.SessionParameter.BindingType] = BindingType.AtomPub;
parameters[DotCMIS.SessionParameter.AtomPubUrl] = contextExecution.UrlGED + "cmisatom";
parameters[DotCMIS.SessionParameter.User] = contextExecution.LoginGED;
parameters[DotCMIS.SessionParameter.Password] = contextExecution.PwdGED;
SessionFactory factory = SessionFactory.NewInstance();
ISession session = factory.GetRepositories(parameters)[0].CreateSession();

// [...]Après la connexion et la déclaration de quelques variables
//On récupère d'id du folder Archives en requête CMIS (répertoire d'Alfresco avec droits limités dans lequel seront déplacés les fichiers envoyés au SAE)
queryGetDoc = "SELECT cmis:objectId FROM cmis:folder WHERE cmis:name='Archives'";
docResults = session.Query(queryGetDoc, false);
searchIdFolder = docResults.FirstOrDefault();
//On lit la valeur de l'id de l'objet folder
idDest = new ObjectId(searchIdFolder["cmis:objectId"].FirstValue.ToString());

//Construction requête CMIS qui va aller chercher toute les métadonnées des documents ayant l'aspect transfertToSAE
String query = "SELECT D.*, M.*, T.*, O.* " +
    "FROM cmis:document as D " +
    "JOIN my:transfertToSAE AS M ON D.cmis:objectId = M.cmis:objectId " +
    "JOIN cm:titled AS T ON D.cmis:objectId = T.cmis:objectId " +
    "JOIN cm:taggable AS O on D.cmis:objectId = O.cmis:objectId ";
//Erreur de servlet lorsque l'on met la condition suivante :
//WHERE M.my:isToArchive = true ;
//Pour l'instant, on récupère donc tous les documents qui ont l'aspect créé de défini (peu de documents l'ont)

//Exécution de la requête CMIS
docResults = session.Query(query, false);

//Pour chaque résultat
foreach (IQueryResult docHit in docResults)
{
    //On récupère l'identifiant CMIS de l'objet
    docId = docHit["D.cmis:objectId"].FirstValue.ToString();

    //On le transforme en document car il s'agit forcément d'un doc
```

¹ CDG56 : Centre de Gestion du 56

```

IDocument document = session.GetObject(docId.ToString()) as IDocument;
Document doc = (Document)document;
//Si le doc est à archiver et fait parti du folder SAE (on ne peut pas ajouter cette
condition en jointure ou autre lors de la 1ère requête)
if (docHit["M.my:isToArchive"].FirstValue != null &&
docHit["M.my:isToArchive"].FirstValue.ToString().ToUpper() == "TRUE")
{
    string entrepotOrigine = doc.Parents[0].Path;
    queryGetDoc = "SELECT cmis:objectId FROM cmis:folder WHERE cmis:name='"
    entrepotOrigine + "'";
    docResults = session.Query(queryGetDoc, false);
    searchIdFolder = docResults.FirstOrDefault();
    //On lit la valeur de l'id de l'objet entrepôt où se trouve le document à archiver
    idDest = new ObjectId(searchIdFolder["cmis:objectId"].FirstValue.ToString());

    //On récupère le contenu du fichier
    Stream strFile = doc.GetContentStream().Stream;
    byte[] compressedBytes;
    string hachSha256 = "";
    SHA256 mySHA256 = SHA256Managed.Create();
    using (MemoryStream ms = new MemoryStream())
    {
        strFile.CopyTo(ms);
        //Si le document est déjà de type zip, on ne le compresse pas
        try
        {
            //On transforme en byte la memory stream
            byte[] fileContent = ms.ToArray();
            //On récupère le hash en SHA256 de notre fichier
            hachSha256 =
                BitConverter.ToString(mySHA256.ComputeHash(fileContent)).Replace("-", String.Empty);
            //On génère un Stream compressé
            using (var outStream = new MemoryStream())
            {
                using (var archive = new ZipArchive(outStream,
                ZipArchiveMode.Create, true))
                {
                    var fileInArchive =
                        archive.CreateEntry(doc.ContentStreamFileName,
                        CompressionLevel.Optimal);
                    using (var entryStream = fileInArchive.Open())
                    using (var fileToCompressStream = new
                    MemoryStream(fileContent))
                    {
                        fileToCompressStream.CopyTo(entryStream);
                    }
                }
                //tableau de bytes, mais compressés
                compressedBytes = outStream.ToArray();
            }
        }
    }
}
[...]
[...] On appelle la fonction qui va générer notre bordereau SEDA

```

[...] Extrait de la génération du bordereau SEDA (Là encore, la fonction n'est pas complète, mais le principe reste le même pour tous les nœuds XML. En fonction du profil d'archives utilisé, il faut en effet adapter les nœuds XML à créer dans le fichier généré)

```

/// <summary>
/// Génération du bordereau de transfert SEDA

```

```

/// </summary>
/// <param name="vChaineConnexion">Chaine de connexion à la BDD postgresql d'as@lae</param>
/// <param name="hashSha256">Chaîne contenant le hash en SHA256 du fichier à archiver</param>
/// <param name="docResult">Document de type IQueryResult à archiver</param>
/// <returns>Bordereau de transfert SEDA</returns>
public static XmlDocument GenerateBordereauXmlSeda_0_2(string vChaineConnexion, string hashSha256, IQueryResult docResult)
{
try {
  XmlDocument doc = new XmlDocument();
  XmlNode docNode = doc.CreateXmlDeclaration("1.0", "UTF-8", null);
  string ns = "fr:gouv:ae:archive:draft:standard_echange_v0.2";
  doc.AppendChild(docNode);

  XmlNode archiveTransferNode = doc.CreateElement("ArchiveTransfer", ns);
  doc.AppendChild(archiveTransferNode);

  XmlNode commentNode = doc.CreateElement("Comment", ns);
  if (docResult["M.my:comment"].FirstValue != null)
  {
    commentNode.AppendChild(doc.CreateTextNode(docResult["M.my:comment"].FirstValue.ToString()));
  }
  archiveTransferNode.AppendChild(commentNode);

  XmlNode dateNode = doc.CreateElement("Date", ns);
  dateNode.AppendChild(doc.CreateTextNode(DateTime.Now.ToString("yyyy-MM-dd") +
"T00:00:00+02:00"));
  archiveTransferNode.AppendChild(dateNode);

  //Identifiant du transfert
  //Ici, on va récupérer la prochaine valeur directement en base de données d'As@lae.
  Aucun web service ne permet de récupérer cette valeur ou d'autres, il s'agit ici uniquement
  d'un test pour vérifier que l'on peut récupérer des informations
  XmlNode transferIdNode = doc.CreateElement("TransferIdentifier", ns);
  string nextId = "MA_" + Asalae.getNextTransferId(vChaineConnexion);
  transferIdNode.AppendChild(doc.CreateTextNode(nextId));
  archiveTransferNode.AppendChild(transferIdNode);

  //Identification du serveur versant, mis en dur, mais devrait être récupéré
  //automatiquement
  XmlNode transferringAgencyNode = doc.CreateElement("TransferringAgency", ns);
  archiveTransferNode.AppendChild(transferringAgencyNode);
  XmlNode transferringAgencyIdNode = doc.CreateElement("Identification", ns);
  transferringAgencyIdNode.AppendChild(doc.CreateTextNode("FRAD001"));
  transferringAgencyNode.AppendChild(transferringAgencyIdNode);

  // [...]Autres déclarations

  //Noeud intégrité
  XmlNode integrityNode = doc.CreateElement("Integrity", ns);
  archiveTransferNode.AppendChild(integrityNode);

  //Hash du fichier en SHA256
  XmlNode containsSha256Node = doc.CreateElement("Contains", ns);
 XmlAttribute urlSha256Attr = doc.CreateAttribute("algorithme");
  urlSha256Attr.Value = "http://www.w3.org/2001/04/xmlenc#sha256";
  containsSha256Node.Attributes.Append(urlSha256Attr);
  containsSha256Node.AppendChild(doc.CreateTextNode(hashSha256));
  integrityNode.AppendChild(containsSha256Node);

  //identifiant de l'unité (nom du fichier)
}

```

```

XmlNode unitIdNode = doc.CreateElement("UnitIdentifier", ns);
unitIdNode.AppendChild(doc.CreateTextNode(docResult["D.cmis:name"].FirstValue.ToString()));
integrityNode.AppendChild(unitIdNode);

// [...] Autres déclarations

// Nom du contenu
XmlNode nameContainNode = doc.CreateElement("Name", ns);
if (docResult["T.cm:title"].FirstValue != null)
{
    nameContainNode.AppendChild(doc.CreateTextNode(docResult["T.cm:title"].FirstValue.ToString()));
}
containsNode.AppendChild(nameContainNode);

// Noeud description du contenu
XmlNode contentDescriptionNode = doc.CreateElement("ContentDescription", ns);
containsNode.AppendChild(contentDescriptionNode);

// Description du contenu
XmlNode descriptionNode = doc.CreateElement("Description", ns);
if (docResult["T.cm:description"].FirstValue != null)
{
    descriptionNode.AppendChild(doc.CreateTextNode(docResult["T.cm:description"].FirstValue.ToString()));
}
contentDescriptionNode.AppendChild(descriptionNode);

// [...] Autres déclarations

// Noeud sort final
XmlNode appraisalNode = doc.CreateElement("Appraisal", ns);
containsNode.AppendChild(appraisalNode);

// Code sort final
XmlNode codeSortNode = doc.CreateElement("Code", ns);
XmlAttribute codeSortAttr = doc.CreateAttribute("listVersionID");
codeSortAttr.Value = "edition 2009";
codeSortNode.Attributes.Append(codeSortAttr);
if (docResult["M.my:sortFinal"].FirstValue != null)
{
    codeSortNode.AppendChild(doc.CreateTextNode(docResult["M.my:sortFinal"].FirstValue.ToString().ToLower()));
}
appraisalNode.AppendChild(codeSortNode);

// Durée sort final
XmlNode dureeSortNode = doc.CreateElement("Duration", ns);
if (docResult["M.my:dureeSortFinal"].FirstValue != null)
{
    dureeSortNode.AppendChild(doc.CreateTextNode("P" +
docResult["M.my:dureeSortFinal"].FirstValue.ToString() + "Y"));
}
appraisalNode.AppendChild(dureeSortNode);

// Date de départ sort final
XmlNode startDateSortNode = doc.CreateElement("StartDate", ns);
if (docResult["M.my:duaDateDeb"].FirstValue != null)
{
    string dateDebDUA =
Convert.ToDateTime((DateTime)(docResult["M.my:duaDateDeb"].FirstValue)).ToString("yyyy-MM-
dd");
    startDateSortNode.AppendChild(doc.CreateTextNode(dateDebDUA));
}
appraisalNode.AppendChild(startDateSortNode);

```

```

//Noeud communicabilité
XmlNode accessRestrictionNode = doc.CreateElement("AccessRestriction", ns);
containsNode.AppendChild(accessRestrictionNode);

//Code communicabilité
XmlNode codeCommNode = doc.CreateElement("Code", ns);
XmlAttribute descCodeCommAttr1 = doc.CreateAttribute("listVersionID");
descCodeCommAttr1.Value = "edition 2009";
codeCommNode.Attributes.Append(descCodeCommAttr1);
if (docResult["M.my:regleComm"].FirstValue != null)
{
    string codeAccessRestriction =
docResult["M.my:regleComm"].FirstValue.ToString().Substring(0,
docResult["M.my:regleComm"].FirstValue.ToString().IndexOf(" "));
    codeCommNode.AppendChild(doc.CreateTextNode(codeAccessRestriction));
}
accessRestrictionNode.AppendChild(codeCommNode);

//Date communicabilité //comDateDeb
XmlNode startDateCommNode = doc.CreateElement("StartDate", ns);
if (docResult["M.my:comDateDeb"].FirstValue != null)
{
    string dateDebComm =
Convert.ToDateTime(DateTime)(docResult["M.my:comDateDeb"].FirstValue).ToString("yyyy-MM-
dd");
    startDateCommNode.AppendChild(doc.CreateTextNode(dateDebComm));
}
accessRestrictionNode.AppendChild(startDateCommNode);

//Noeud document
XmlNode documentNode = doc.CreateElement("Document", ns);
containsNode.AppendChild(documentNode);

//Nom du document joint
XmlNode attachmentNode = doc.CreateElement("Attachment", ns);
XmlAttribute filenameAttr = doc.CreateAttribute("filename");
filenameAttr.Value = docResult["D.cmis:name"].FirstValue.ToString();
attachmentNode.Attributes.Append(filenameAttr);
documentNode.AppendChild(attachmentNode);

XmlNode purposeNode = doc.CreateElement("Purpose", ns);
if (docResult["M.my:object"].FirstValue != null)
{
    purposeNode.AppendChild(doc.CreateTextNode(docResult["M.my:object"].FirstValue.ToString()));
}
documentNode.AppendChild(purposeNode);

// [...] Autres déclarations

doc.Save(Console.Out);
return doc;
// [...] Fin fonction avec gestion erreurs

[...] Extrait de la méthode d'appel du web service REST d'As@iae (Méthode adaptée à
partir du code source fourni par Patrick Percot du CDG56)
/// <summary>
/// Versement d'une archive à as@iae via REST
/// </summary>
/// <param name="bordereauStream">MemoryStream du bordereau de transfert SEDA</param>
/// <param name="fileNameToArchive">Nom du fichier à archiver</param>
/// <param name="contentArchive">byte[] du fichier à archiver</param>

```

```

/// <param name="uriRest">Uri du service REST</param>
/// <returns>Message de retour du SAE</returns>
public string ArchiveZipFileToAsalae(MemoryStream bordereauStream, string
FileNameToArchive, byte[] contentArchive, string accuseMailAdress, string uriRest)
{
    string boundary = "-----" + DateTime.Now.Ticks.ToString("x");
    FileNameToArchive = FileNameToArchive.Substring(0, FileNameToArchive.LastIndexOf('.')) +
    ".zip";
    using (MemoryStream fileContentStream = new MemoryStream())
    {
        using (Stream mailContentStream = new MemoryStream())
        {
            fileContentStream.Write(contentArchive, 0, contentArchive.Length);
            try
            {
                ZipArchive archive;
                //Vérifie que le fichier archive est bien formé
                archive = new ZipArchive(new MemoryStream(contentArchive),
ZipArchiveMode.Read);
            }
            catch (InvalidDataException e)
            {
                string retError = "Le fichier n'est pas un ZIP ou est corrompu : " +
e.Message;
                if (traceActions)
                    tracesWriter.WriteLine(retError);
                throw new Exception(retError);
            }

            //send
            using (var content = new MultipartFormDataContent(boundary))
            {
                bordereauStream.Seek(0, SeekOrigin.Begin);
                HttpContent bordereauHttpContent = new StreamContent(bordereauStream);
                bordereauHttpContent.Headers.ContentType =
MediaTypeHeaderValue.Parse("text/xml");
                content.Add(bordereauHttpContent, "\"seda_message\"",
HttpUtility.UrlEncode("bordereau.xml"));

                fileContentStream.Seek(0, SeekOrigin.Begin);
                HttpContent fileHttpContent = new StreamContent(fileContentStream);
                fileHttpContent.Headers.ContentType =
MediaTypeHeaderValue.Parse("application/x-zip-compressed");
                content.Add(fileHttpContent, "\"attachments\"",
HttpUtility.UrlEncode(Path.GetFileName(FileNameToArchive)));

                //Appel du webservice rest
                Task<HttpResponseMessage> response = this.RunAsync(uriRest,
HttpMethod.Post, content);
                bool isOK = response.Result.IsSuccessStatusCode;
                string resultat = response.Result.Content.ReadAsStringAsync().Result;

                if (response.Result.IsSuccessStatusCode)
                {
                    return "OK;" + response.Result.Content.ReadAsStringAsync().Result;
                }
                else
                {
                    return "KO; Une erreur s'est produite lors de l'appel REST " +
response.Result.Content.ReadAsStringAsync().Result;
                }
            }
        }
    }
}
// [...]Fin fonction et gestion erreurs
  
```

**MÉMOIRE D'INGÉNIEUR DU CNAM EN INFORMATIQUE,
OPTION ARCHITECTURE ET INGÉNIERIE DES SYSTÈMES ET LOGICIELS**

**Étude et réalisation d'un prototype d'archivage électronique appliquée aux
collectivités territoriales**

Julien MONTEL

Rennes 2015

Résumé

La dématérialisation des documents fait que l'archivage électronique est devenu incontournable. Rennes Métropole et la Ville de Rennes se doivent de réfléchir à la mise en place d'un système d'archivage électronique (SAE) définitif. Outre les documents ayant une durée d'utilité administrative, les collectivités territoriales génèrent également des documents ayant une valeur historique, tels que les marchés publics, les documents d'état civil ou encore la cartographie des réseaux d'énergie.

Ce mémoire est rédigé afin de faciliter la spécification et le choix d'une telle solution. Il présente un état de l'art à travers les contextes législatif, normatif et technologique en vigueur dans le domaine. Il traite également de la réalisation et de la mise en œuvre d'un prototype de SAE patrimonial.

Mots-clés: SAE, OAIS, Métadonnées, Archivage Électronique, Formats, Supports, Conservation, Patrimoine

Abstract

Electronic Record Management has become unavoidable today due to document dematerialization. Rennes Métropole and Ville de Rennes need to consider the set-up of a permanent electronic record management system (ERMS). Many documents have administrative usefulness in the long-run, and local authorities produce documents of historical value, such as state contracts, civil registry and energy network maps.

This report aims at facilitating the specification and choice for such a solution, with reference to the existing legislative, normative and technological framework in the area of document dematerialization. It also deals with the creation and implementation of a patrimonial ERMS prototype.

Keywords: ERMS, OAIS, Metadata, Electronic Record Management, Support platforms, Media, Preservation, Heritage