



UFR SHA

Mention Information-Communication
Spécialité Documentation

Année universitaire 2016-2017

**Le potentiel des technologies du web sémantique pour
l'archivage dans un contexte de gouvernance de l'information :
*le cas du Système d'Archivage Électronique Mutualisé (SAEM) girondin***

Mémoire pour l'obtention du Master esDOC

Présenté par

Alexander Felton

Le 22 septembre 2017

Sous la direction de

Monsieur David Guillemin
Université de Poitiers





UFR SHA

Mention Information-Communication
Spécialité Documentation

Année universitaire 2016-2017

**Le potentiel des technologies du web sémantique pour
l'archivage dans un contexte de gouvernance de l'information :
*le cas du Système d'Archivage Électronique Mutualisé (SAEM) girondin***

Mémoire pour l'obtention du Master esDOC

Présenté par

Alexander Felton

Le 22 septembre 2017

Sous la direction de

Monsieur David Guillemain
Université de Poitiers



Remerciements

Je tiens tout d'abord à remercier Monsieur David Guillemin, mon directeur de mémoire, pour son écoute, sa patience et ses nombreux conseils.

Mes remerciements vont également à l'ensemble des personnes avec qui j'ai eu la chance de travailler lors de mon stage au Conseil départemental de la Gironde, et plus particulièrement à Pascal Romain et à Delphine Jamet, pour leur accueil et leur disponibilité.

Je remercie l'ensemble des professionnels qui ont accepté de répondre à mes questions, me permettant de réunir les matériaux nécessaires pour mener à bien mon travail.

Enfin, je tiens à remercier mes proches, qui m'ont soutenu et encouragé tout au long de la rédaction de ce mémoire.

Table des abréviations

AD : Archives Départementales

ARK : Archival Resource Key

EAD : Encoded Archival Description

EGAD : Groupe d'experts sur la description archivistique

GED : Gestion Électronique des Documents

GI : Gouvernance de l'Information

LOD : Linked Open Data

OAIS : Open Archival Information System

OWL : Ontology Web Language

RDF : Resource Description Framework

RiC : Records in Context

SAE : Système d'Archivage Électronique

SAEM : Système d'Archivage Électronique Mutualisé

SEDA : Standard d'Échange de Données pour l'Archivage

SI : Système d'Information

SIA : Système d'Information Archivistique

SIAF : Service Interministériel des Archives de France

SKOS : Simple Knowledge Organisation System

URI : Uniform Resource Identifier

W3C : World Wide Web Consortium

XML : Extensible Markup Language

Sommaire

Remerciements

Table des abréviations

Sommaire

Introduction

Première partie : l'impact du numérique et des évolutions du web sur le traitement des données archivistiques

- I. L'impact du numérique sur la pratique de l'archivage
- II. De l'élaboration de modèles conceptuels à la mise en place de référentiels : les enjeux d'une gouvernance de l'information et des données
- III. Optimiser la gouvernance dans un service d'archives : les potentialités des technologies du web sémantique

Deuxième partie : l'élaboration du protocole d'expérimentation

- I. L'étude de cas à partir de l'exemple du SAEM girondin : attentes et objectifs
- II. L'approche qualitative pour compléter et approfondir les résultats
- III. Procédure d'analyse et d'interprétation des résultats

Troisième partie : résultats de l'expérimentation

- I. La transition numérique des services d'archives, symbole d'une bascule dans les pratiques professionnelles
- II. Le nouveau paradigme de la gouvernance de l'information : vers davantage de transversalité
- III. Web sémantique et données archivistiques, un enchevêtrement souhaité et souhaitable ?

Conclusion

Table des matières

Bibliographie

Introduction

Depuis plus de deux siècles, en France, les services publics d'archives assurent la sauvegarde d'un précieux patrimoine écrit et sont considérés comme les gardiens de la mémoire du pays. Les archives présentent une double nature, puisqu'elles sont à la fois les composantes d'un patrimoine et de véritables outils administratifs, dans la mesure où ces documents permettent le bon fonctionnement des institutions. La révolution numérique redéfinit ce champ d'action, introduisant des ambiguïtés par rapport à la notion même de document, mais également en bouleversant le rapport du citoyen à l'information. Fruit d'une longue bataille menée par la communauté des archivistes français, notamment par l'AAF¹, l'intégration de la notion de données, qui vient alors se juxtaposer au terme de « documents », à la définition des archives à l'occasion de la promulgation de la loi LCAP², cristallise cette transition numérique.

C'est alors toute la profession, dont la mission est, depuis son apparition, la consignation d'écrits authentiques et fiables, qui voit ses procédures modifiées en profondeur par ce changement d'objet. La numérisation des fonds patrimoniaux et la mise en ligne massive de documents sur les interfaces de consultation des services d'archives, accompagnée par une dématérialisation d'une grande partie de la production, constituent le point de départ d'une mutation des pratiques des archivistes. Dès lors, c'est le support même avec lequel les archivistes ont toujours travaillé qui est remis en question, fragilisé par un environnement caractérisé par des obsolescences technologiques à courte échéance et des impératifs de traitement et de conservation plus rigoureux que jamais. D'autre part, c'est la diversité même des acteurs concernés par la pratique de l'archivage, en raison des considérations technologiques et juridiques introduites par cette dématérialisation des processus, qui subit d'importantes transformations depuis la genèse de l'administration électronique, à la fin des années 1990.

Ces bouleversements ne sont pour autant pas propres aux services d'archives, puisque c'est la place de l'information, et par extension de la donnée, qui fait l'objet d'une

1 Association des archivistes français

2 Loi relative à la liberté de la création, à l'architecture et au patrimoine, promulguée en juillet 2016.

véritable réévaluation au sein même des organisations, et qui se voit attribuée, à l'heure du big data³ et de l'open data⁴, une valeur stratégique sans précédent. Ce sont de nouveaux modes d'organisation qui émergent alors, avec une multiplication des acteurs qui doivent apprendre à composer avec la donnée. Ce nouveau paradigme s'est vu attribué, depuis le tournant des années 2010, le titre de gouvernance de l'information. Dans son rapport adressé à la ministre de la Culture en mars dernier, Christine Nougaret souligne ces enjeux en matière de gouvernance pour les services d'archives :

« Pour les archives, l'irruption du numérique est indubitablement une chance qui nous est donnée de gagner une visibilité accrue dans nos sociétés et de convaincre nos interlocuteurs que les archives sont les outils d'une meilleure gouvernance dans les organismes. C'est seulement si nous gagnons ce pari que nous pourrions préserver ce nouveau patrimoine qui se construit et qui, pour l'heure, est encore infiniment fragile. »⁵

En raison de ces mutations, ce sont les compétences des archivistes qui doivent évoluer, avec une réelle prise en compte de l'archivage numérique, dans la mesure où celui-ci revêt désormais un réel intérêt stratégique pour toute institution. Ce sont alors, comme le souligne là aussi Christine Nougaret, les référentiels métiers qui doivent intégrer cette dimension numérique, qui est désormais incontournable pour tout archiviste qui souhaite participer à des projets innovants. L'ensemble du paysage de la collecte des archives, mission essentielle de ces services, se voit alors bouleversé par l'explosion documentaire propre au numérique et doit par conséquent s'appuyer sur des technologies adaptées et des processus métiers harmonisés, sans pour autant être à la source d'une dichotomie entre l'archivage papier et l'archivage numérique. Parallèlement, puisque c'est l'accès aux archives qui justifie avant tout la collecte, et que la valorisation et la communication de ce patrimoine figurent parmi les missions historiques des archivistes, les évolutions liées à l'apparition et au développement du web contribuent à ce changement de paradigme. En effet, parallèlement à l'informatisation des services et à la numérisation massive des fonds d'archives à partir des années 1990, ce sont les habitudes de consultation des usagers qui ont connu d'importantes transformations, puisque les salles de lecture ont été détrônées par les

3 Peut être traduit par données massives. Correspond à l'explosion, en termes de volumétrie, des données numériques.

4 Mouvement pour l'ouverture des données publiques.

5 NOUGARET, Christine. Une stratégie nationale pour la collecte et l'accès aux archives à l'ère numérique. *FranceArchives* [en ligne]. [Consulté le 31 juillet 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/28204701>

sites web et interfaces de consultation en ligne. Ces dernières ont elles aussi connu leur lot d'évolutions techniques et ergonomiques, à la faveur, dans un premier temps, du développement de la dimension contributive du web, avec une implication des usagers dans les pratiques d'indexation et de description des fonds. L'avènement de la donnée est également l'occasion d'une transformation du web, puisque semble émerger aujourd'hui un web « 3.0 », ou web des données, reposant sur des technologies permettant de créer du sens et de mettre en relation des données d'horizons divers. Le rapport Quenet préconisait déjà, en 2011, la mise en œuvre d'un portail unifié d'accès aux archives devant « exploiter les potentialités du web sémantique et développer des outils innovants de navigation entre les fonds et adaptés aux nécessités de la recherche. »⁶ Maurice Quenet y mentionnait le besoin d'interopérabilité⁷ des choix techniques, évoquant notamment la problématique de la pérennité des liens vers les ressources en ligne. Pour autant, la prise en compte des enjeux liés à un web des données, encore émergent, n'y est que succinctement évoquée et ce n'est qu'en 2014 que le ministère de la Culture s'empare de ces questions, à travers la publication d'une feuille de route stratégique, dans laquelle est reconnu le potentiel du web sémantique pour les données culturelles :

« La souplesse et la puissance des technologies du Web sémantique dites technologies 3.0 transforment dans ces conditions le Web en une immense base de connaissance ouverte en perpétuelle évolution, et offrent ainsi des perspectives de développement stratégiques pour les institutions culturelles. »⁸

En raison de ce contexte protéiforme et des particularités des archives mais également du fonctionnement des services publics qui sont responsables de leur traitement comme de leur conservation, dans quelle mesure les données issues du traitement des archives publiques peuvent-elles bénéficier des apports des technologies du web sémantique ? S'agit-il avant tout d'harmoniser les procédures de traitement et de conservation, via un recours à des référentiels communs et à des identifiants uniques

6 QUENET, Maurice. Quel avenir pour les Archives de France ? [en ligne]. 2011. [Consulté le 28 août 2017]. Disponible à l'adresse : <http://www.ladocumentationfrancaise.fr/rapports-publics/114000194/index.shtml>

7 « L'interopérabilité est la capacité que possède un produit ou un système, dont les interfaces sont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre. » Cf. <http://definition-interoperabilite.info/>

8 Feuille de route stratégique – Métadonnées culturelles et transition Web 3.0 [en ligne]. *Ministère de la Culture et de la Communication*, 2014. [Consulté le 27 août 2017]. Disponible à l'adresse : http://www.culturecommunication.gouv.fr/var/culture/storage/pub/feuille_de_route__metadonnees_culturelles_et_transition_web_3_0_janvier_2014/index.htm

permettant de satisfaire les besoins de pérennité des documents archivés dans un environnement informatisé ? Ou alors s'agit-il véritablement d'envisager une mise en relation avec l'ensemble des données du secteur culturel, afin de contribuer à la diffusion du patrimoine et de la culture par le biais de l'interconnexion que permet le web des données liées ?

Si la mise en œuvre de dispositifs techniques garantissant une interopérabilité entre les pratiques des archivistes et les standards introduits par le web des données est cruciale, et que nous ne pourrions faire l'économie de son analyse, nous tâcherons d'aller au-delà de ce cadre et de nous intéresser aux enjeux professionnels qui accompagnent ces mutations, en nous penchant sur la manière dont ces évolutions sont perçues par les acteurs concernés. La première partie de ce mémoire sera alors dédiée à un état de l'art de la littérature concernant l'ensemble des aspects qui sous-tendent notre problématique. Nous nous intéresserons alors tout d'abord à l'impact du numérique sur la pratique de l'archivage, avant d'étudier le contexte de la gouvernance de l'information dans lequel doivent évoluer les archivistes, puis nous ciblerons la manière dont les technologies du web sémantique peuvent rejoindre les problématiques qui caractérisent l'archivage à l'ère du numérique, en étudiant notamment les caractéristiques du web des données. Dans la seconde partie, nous détaillerons la manière dont nous avons procédé pour mettre au point un protocole d'expérimentation permettant d'analyser de manière plus concrète les différentes problématiques préalablement établies. Nous y présenterons notamment le contexte dans lequel cette expérimentation a pris place, ainsi que la méthodologie adoptée. Nous partirons d'un exemple concret, celui du système d'archivage électronique mutualisé girondin, dans lequel les technologies du web sémantique ont fait l'objet d'une attention particulière. Enfin, nous mettrons en avant les résultats obtenus, dans l'optique d'une confrontation avec la théorie et en vue de vérifier la validité des différentes hypothèses que nous aurons pu émettre, tout en essayant, dans la mesure du possible, d'élargir nos observations à l'ensemble du paysage des archives publiques en France à l'ère du web 3.0.

Première partie : l'impact du numérique et des évolutions du web sur le traitement des données archivistes

L'objectif principal de cette première partie sera de contextualiser le sujet, en explicitant les différentes notions qui permettent de le délimiter, avant de s'intéresser à la manière dont se complètent les différentes couches de notre objet d'étude. Il s'agit donc en premier lieu de replacer le sujet dans un contexte plus global, celui de l'évolution des technologies du numérique, qui bouleverse aussi bien les pratiques des archivistes que celles des bibliothécaires et documentalistes. Par ailleurs, les enjeux qui en découlent transcendent la problématique des pratiques des professionnels de l'information, puisque ce sont de véritables bouleversements des usages qui semblent se profiler avec l'avènement d'un web dit « 3.0 ». La question de l'appropriation des technologies semble alors se poser de manière transversale au sein des organisations et il s'agira d'étudier l'impact de ces multiples transformations sur les processus métiers des archivistes, à travers l'exemple particulièrement parlant de l'archivage électronique. Cette étape de contextualisation du sujet, basée sur des ressources d'horizons divers, notamment issues de la littérature professionnelle, permettra alors d'explicitier le cheminement qui mènera à la problématisation de notre travail et à la formulation d'hypothèses qui devront être vérifiées par le biais d'une enquête de terrain.

I. L'impact du numérique sur la pratique de l'archivage

I.1. Les archives, produit documentaire de l'activité humaine

Le document d'archives, éléments de définition

Afin d'être en mesure d'aborder l'impact de l'évolution du numérique et du web sur le traitement des archives, il convient d'étudier ce qui fait la spécificité de ces documents et d'analyser en quoi ceux-ci justifient une approche particulière, non seulement en raison de leur nature mais également au regard de l'historique des pratiques professionnelles qui leur sont propres. Ainsi, il s'agit dans un premier temps de définir

la notion de document d'archives afin de saisir ce qui caractérise cet objet d'étude. Les archives ont fait l'objet d'une définition officielle et juridique pour la première fois en France dans la loi du 3 janvier 1979 : « les archives sont l'ensemble des documents, quels que soient leur date, leur forme et leur support matériel, produits et reçus par toute personne physique ou morale, et par tout service ou organisme public ou privé, dans l'exercice de leur activité. »⁹ Leur rôle de témoignage de l'histoire n'est pas alors leur portée première, puisqu'ils sont avant tout nécessaires au fonctionnement des services qui les produisent, notamment en tant que preuve de leur activité.¹⁰ Les archives disposent d'un véritable caractère organique, puisqu'elles découlent d'une activité précise et répondent à des besoins fonctionnels et de gestion. Ces documents font donc l'objet d'une double nature puisqu'ils « constituent des preuves pour leurs producteurs comme pour leurs bénéficiaires, en même temps que des sources pour la recherche historique. »¹¹ Marie-Anne Chabin, fondatrice du cabinet de conseil Archive 17 et spécialiste de l'archivage électronique, évoque quant à elle, la double fonction du document d'archives en précisant qu'il s'agit d'un « document de preuve et de mémoire »¹², plaçant ces deux enjeux sur un pied d'égalité. La définition du dictionnaire des Archives, publié par l'AFNOR¹³, précise ces éléments en mentionnant que le document d'archives a « par lui-même ou par son support » une valeur probatoire. Cette notion est essentielle pour comprendre que le document d'archives n'est pas un document comme les autres. Les archives ne sont pas uniquement des documents conservés pour leur valeur patrimoniale et qui n'intéresseraient alors que les historiens, mais sont conçues dans l'optique d'une utilité immédiate. Ainsi, elles existent dès leur création, et pas uniquement à partir du moment où elles sont versées à un service d'archives, et donc archivées.

9 Code du patrimoine – Article L211-1.

10 En effet, d'après cette définition, désormais intégrée au Code du patrimoine, les archives « ne sont pas seulement – comme on le pense encore trop souvent – les documents anciens qui ne présentent plus d'utilité que pour les historiens ou les curieux. » GALLAND, Bruno, 2016. Les archives. In : *Que sais-je ?* 27 décembre 2016. p. 3-6.

11 Ibid.

12 CHABIN, Marie-Anne. Qu'est-ce qu'un document d'archives ? *TRANSARCHIVISTIQUE* [en ligne]. 2017. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://transarchivistique.fr/quest-ce-quun-document-darchives/>

13 ASSOCIATION FRANÇAISE DE NORMALISATION et ÉCOLE NATIONALE DES CHARTES. Dictionnaire des archives : français-anglais-allemand : de l'archivage aux systèmes d'information. AFNOR, 1991. ISBN 978-2-12-484221-6.

La majeure partie de leur valeur découle de leur contexte de production et c'est la provenance de ces documents qui permet avant tout de les caractériser et d'en retirer de l'information. En raison de cette nature particulière, le document d'archives doit posséder certaines qualités afin d'être considéré comme ce vecteur de la mémoire que présente le Conseil international des archives¹⁴. Celui-ci, pour être un document « digne de confiance », se doit de remplir des critères d'authenticité, de fiabilité, d'intégrité et d'accès. Pour répondre à ces différents critères qui caractérisent la valeur d'un document d'archives, les archivistes ont adapté leurs pratiques, jugeant qu'il était « fondamental de maintenir ensemble les documents de même origine et de ne pas les mélanger avec d'autres de provenance différente [...] car c'est seulement si l'on connaît bien l'origine du document que l'on peut vérifier son authenticité et comprendre la manière dont telle ou telle affaire a pu être traitée »¹⁵. C'est ce que les archivistes appellent, depuis le XIXe siècle, le principe de respect des fonds.

Les fondements de l'archivage : les principes de respect du fonds et la description à plusieurs niveaux

Le document d'archives est effectivement sujet à une interprétation qui reste intimement liée à son contexte de production et à sa provenance. Ces notions se retrouvent dans la pratique de l'archivage puisqu'un document d'archives (c'est-à-dire une « pièce » dans le vocabulaire archivistique) n'est jamais conservé seul mais doit être systématiquement rattaché à un dossier, dans lequel il trouve tout son sens.¹⁶ Un fonds d'archives est alors défini comme « l'ensemble de documents de toute nature constitué de façon organique par un producteur dans l'exercice de ses activités et en fonction de ses attributions. »¹⁷ La pratique de la description archivistique, qui vise à restituer le contenu des fonds, permet d'explicitier cette conception particulière du document, puisqu'elle répond à un impératif qui découle directement de ces

14 INTERNATIONAL COUNCIL ON ARCHIVES, *Que sont les archives ?* [en ligne]. [Consulté le 24 mai 2017]. Disponible à l'adresse : <http://www.ica.org/fr/quest-ce-que-les-archives>.

15 GALLAND, Bruno. Chapitre III. Les missions des archivistes. *Que sais-je ?*, 27 décembre 2016, pp. 85-120.

16 C'est également pour cette raison qu'un simple classement thématique ne peut être mis en application en archivistique puisqu'il favoriserait un aspect plutôt qu'un autre et ôterait une partie de son sens au document.

17 Ibid. Cette notion symbolise la particularité des archives par rapport aux autres documents, puisqu'elle est en réelle opposition avec celle de collection, définie comme la « réunion artificielle de documents en fonction de critères communs liés à leur contenu ou à leur support et dont la juxtaposition est le fruit de la volonté ou du hasard » et plus prégnante dans les bibliothèques.

considérations : la description à plusieurs niveaux. La description des documents d'archives repose entièrement sur ces concepts et s'effectue en fonction de ces différents niveaux dits « organiques », du fonds à la pièce, en passant par les séries et sous-séries organiques, ainsi que les dossiers.

À travers ces principes, qui sont encore aujourd'hui fondamentaux en archivistique, c'est l'importance de la provenance du document qui est mise en lumière. Pour autant, Céline Guyon, vice-présidente de l'association des archivistes français, en s'appuyant sur la vision de Michel Duchein, ancien inspecteur des Archives nationales, rappelle que cette notion de provenance ne doit pas nécessairement être considérée comme un état de fait, puisque « en archivistique contemporaine, cette notion est devenue extrêmement imprécise et complexe. »¹⁸ D'après elle, c'est notamment le numérique qui risque de bouleverser cette notion qui jusque-là a fait l'objet d'une application rigoureuse, et que la manière dont les versements d'archives sont effectués, puisqu'ils sont désormais soumis à une logique de flux, doit faire l'objet d'une remise en question.

Cycle de vie et processus de collecte, quelle application dans l'environnement numérique ?

Le concept de cycle de vie, appliqué largement à l'ensemble des ressources documentaires, trouve également une signification particulière lorsqu'il est appliqué à des documents d'archives, puisque ceux-ci ont fait l'objet, notamment depuis la seconde moitié du XXe siècle, d'une tentative de rationalisation visant à faciliter leur traitement en fonction de leur « parcours de vie »¹⁹. La notion de cycle de vie, en archivistique, découle en effet des besoins d'évaluation qui sont apparus, comme le rappelle Bruno Galland, « lorsqu'on a pris conscience que la masse des documents produits dépassait très largement les capacités de conservation. »²⁰ Dans la tradition archivistique française, c'est la théorie des trois âges, originaire des travaux de l'archiviste Théodore Schellenberg, qui a su trouver un écho favorable et s'imposer comme un standard de fait dès les années 1960. Comme son nom l'indique, le concept

18 GUYON, Céline. La pratique archivistique publique en France, entre adaptation et négociation. *Les Cahiers du numérique*, 17 juin 2015, Vol. 11, n° 2, pp. 77-114.

19 KERN, Gilliane, HOLGADO, Sandra et COTTIN, Michel. Cinquante nuances de cycle de vie. *Les Cahiers du numérique*, 17 juin 2015, Vol. 11, n° 2, pp. 37-76.

20 GALLAND. op. cit. p. 13

repose sur une succession de trois « statuts d'information »²¹, pour reprendre l'expression de Marie-Anne Chabin, permettant de définir le traitement (ou le « sort final ») réservé aux documents concernés à l'issue de leur durée d'utilité administrative (DUA)²². L'essor de la théorie des trois âges a également abouti au développement de nombreux outils, notamment les tableaux de gestion et les calendriers de conservation, qui sont désormais au cœur des pratiques des archivistes. Malgré les critiques émises à de nombreuses reprises²³ par une partie de la communauté des archivistes et records managers, notamment au regard du fait que cette vision « ne permet aucun contrôle des flux documentaires produits dans les systèmes informatiques, où la question de la gestion des espaces de conservation et de l'accès à l'information se pose différemment »²⁴, cette dernière s'est largement imposée et a par ailleurs depuis acquis une légitimité réglementaire, puisqu'elle a été intégrée au Code du patrimoine.

Au final, c'est l'ensemble du processus de collecte qui risque de se voir bouleversé par le numérique. Comme le rappelle Céline Guyon, « le moment de la collecte correspond à la fin de la durée d'utilité administrative. Les documents sont collectés quand ils ne sont plus utiles aux services qui les ont produits pour l'accomplissement de leurs tâches et la justification de leurs décisions »²⁵ Or, pour répondre aux différents critères auxquels doit souscrire un document à valeur probante, le numérique impose d'envisager la prise en charge « dans un système approprié, dès sa validation [...] et non de manière différée. » Ainsi, la collecte et la diffusion des documents se feraient de manière parallèle, puisque « la ligne horizontale du temps généralement utilisée pour représenter le cycle de vie des archives et le passage successif des archives par trois états – archives courantes, intermédiaires et définitives – est brisée. »²⁶ Il s'agira alors à travers notre étude d'interroger la pertinence de cette conception du cycle de vie, qui semble avoir camouflé en France d'autres visions, issues notamment de la pratique du

21 CHABIN. art. cit. p. 11

22 Cf. schéma en annexe 1.

23 GALLAND. op. cit. p. 13. En effet, cette théorie a fait l'objet de nombreuses contestations au fil des années et avec l'émergence de nouvelles façons de concevoir le cycle de vie : « dans de nombreux cas, le cycle de vie du document ne connaît que deux temps, correspondant à sa valeur « primaire » (utilité pour l'organisme qui l'a produit) et à sa valeur « secondaire » (généralement historique). »

24 KERN, HOLGADO et COTTIN. art. cit. p. 13

25 GUYON. art. cit. p. 13

26 Ibid.

records management²⁷, dans le contexte de la mise en place de systèmes d'archivage électroniques.

I.2. De ISAD(G) à EAD : l'évolution des normes de description archivistique

Le quatuor normatif de la description archivistique

Les concepts évoqués précédemment, notamment l'importance de la provenance et du contexte de production, sous-tendent la manière dont les archivistes procèdent pour décrire et indexer les archives. Si ces procédés sont le fruit d'une longue tradition de description archivistique, les premières normes en la matière ont véritablement vu le jour au début des années 1990, sous l'impulsion du Conseil International des Archives (ICA). Ainsi, par rapport aux bibliothécaires, qui dès les années 60 et 70 ont tenté de normaliser la description bibliographique, une telle orientation n'a été envisagée que tardivement par les archivistes.²⁸ Il est alors crucial de nous pencher sur la manière dont a été élaboré ce corpus normatif et l'impact qu'il a sur les pratiques du métier d'archiviste, afin d'être en mesure de nous interroger sur leur interopérabilité avec les dernières évolutions du web. Les quatre normes élaborées par l'ICA sont les suivantes :

- ISAD(G) : norme générale et internationale de description archivistique, qui a pour objet la description des documents d'archives (1994).
- ISAAR(CPF) : norme internationale sur les notices d'autorité archivistiques relatives aux collectivités, aux personnes et aux familles. Elle encadre la réalisation de notices d'autorité décrivant les entités (collectivités, personnes ou familles) associées à la production et à la gestion des archives (1996).
- ISDF : norme internationale pour la description des fonctions. Cette norme a

27 Gestion des documents d'activité, conception d'origine anglo-saxonne du traitement des archives courantes et intermédiaires.

28 GRIMOÛARD, Claire Sibille-de, 2012. Les normes internationales de description archivistique : origines, développements, perspectives. In : *La Gazette des archives*. 2012. Vol. 228, n° 4, p. 77-90. DOI 10.3406/gazar.2012.4985. À cette époque, l'élan de la normalisation a été perçu comme une véritable nouveauté au sein des services d'archives, puisque comme le rappelle Claire Sibille de Grimoüard, « jusque-là, une opinion couramment répandue chez les archivistes était que chaque fonds avait son propre classement, ses caractéristiques propres et qu'il était le résultat d'un processus historique unique [...] On pensait généralement qu'il était inutile d'élaborer des normes pour la description des archives. »

quant à elle pour objet la description des fonctions des collectivités associées à la production et à la gestion des archives.

- ISDIAH : norme internationale pour la description des institutions de conservation des archives.

Dès son introduction, la norme ISAD(G) explicite les besoins auxquels répond la pratique, dans son ensemble :

« L'objet de la description archivistique est d'identifier et d'expliquer le contexte et le contenu des documents d'archives, en vue de faciliter leur accès. »²⁹

Il s'agit donc à la fois pour les professionnels d'être en mesure de décrire, en précisant les caractéristiques des documents collectés et conservés, mais également de préparer le terrain pour la consultation par les usagers. Ainsi, si ces normes se sont appuyées sur les pratiques antérieures des archivistes et respectent les grands principes de la description à plusieurs niveaux cités précédemment, elles constituent avant tout un support de travail et un outil fondamental pour les archivistes. Ce sont elles qui ont permis d'uniformiser les pratiques au niveau international. En effet, ISAD(G) précise les différents éléments (au nombre de 26) à renseigner pour décrire un document d'archives dans sa globalité, en respectant les différents niveaux de classement.

Si ISAD(G) met clairement en lumière la volonté d'une mise en correspondance avec les pratiques de description à plusieurs niveaux et avec la rédaction d'instruments de recherche, la norme ISAAR(CPF) est le signe avant-coureur d'une nouvelle manière de concevoir la description archivistique, puisqu'elle « a introduit dans les pratiques archivistiques l'élaboration de descriptions distinctes et liées des archives et de leur contexte de production, ainsi que le contrôle d'autorité des producteurs, permettant la gestion des relations entre des fonds et plusieurs collectivités, personnes ou familles. »³⁰ Cette norme constitue le premier pas vers une « conception multidimensionnelle » de la description du contexte de production. En effet, elle permet de décrire de manière autonome les producteurs à travers la rédaction de notices d'autorité distinctes, qui vont permettre d'établir le lien entre un producteur et

29 ISAD(G) : Norme générale et internationale de description archivistique – Deuxième édition. *International Council on Archives* [en ligne]. [Consulté le 12 avril 2017]. Disponible à l'adresse : <https://www.ica.org/fr/isadg-norme-generale-et-internationale-de-description-archivistique-deuxieme-edition>

30 GRIMOÛARD, art. cit. p. 15

les séries ou dossiers qu'il produit, sans pour autant les inscrire dans une logique de fonds. Cette façon de procéder facilite alors le partage entre institutions et encourage la communication des notices d'autorité.

Dans la mesure où « les normes archivistiques incarnent les principes fondamentaux de la discipline et facilitent les pratiques de la communauté professionnelle »³¹, il s'agira pour nous d'étudier la manière dont leur transposition à l'ère du numérique a été envisagée par les archivistes et d'analyser en quoi elles ont dû être adaptées pour répondre aux exigences de l'environnement informatique.

Description archivistique et métadonnées, un enchevêtrement naturel ?

Les métadonnées sont, comme leur nom l'indique, des données décrivant d'autres données. Elles permettent alors de décrire et de donner des informations sur le contenu d'un document, indépendamment de son support. Avec l'essor du numérique, le recours à ce type de données répond à la nécessité de garantir la restitution du contenu d'un document et l'exploitabilité de ce dernier. La mise en œuvre de ces métadonnées passe notamment par le recours à un langage structuré. Le XML, qui est un langage ouvert permettant de créer ses propres balises et qui a pour objet la description d'un contenu, s'est imposé comme un standard international pour l'échange de métadonnées.³² À partir de ce langage structuré, des schémas de métadonnées, définissant la structure d'un document XML, vont pouvoir être élaborés pour répondre à des besoins spécifiques.

Il semble alors difficile d'évoquer les normes de description archivistique sans évoquer l'apparition, en parallèle, de schémas de métadonnées qui en découlent directement ou du moins qui ont été pensés pour être interopérables avec ce corpus normatif. Ainsi, le schéma de métadonnées EAD qui a été conçu sous la forme d'une DTD³³ XML par la Société des archivistes américains dès 1998, est représentatif de cette transition numérique de la description archivistique. Si l'EAD n'a pas été conçu au même moment

31 GRIMOÛARD, Claire Sibille-de. Élaborer des normes de description... et les confronter à la pratique d'aujourd'hui. *La Gazette des archives*, 2012, Vol. 226, n° 2, pp. 165-177. DOI 10.3406/gazar.2012.4906.

32 Cf. Ibid. Un document XML correspond à une arborescence d'éléments qui va permettre de décomposer une ressource en unités d'informations et c'est « l'indépendance de la structure d'un document XML par rapport à sa présentation [qui] le rend pérenne. »

33 Document Type Definition, à savoir un fichier XML décrivant les règles structurelles et syntaxiques d'un document XML.

que la norme ISAD(G), il est pensé pour être complètement interopérable avec cette dernière, puisqu'il « combine toutes les composantes de la description archivistique (les archives et leur contexte) » et fournit un ensemble de balises (146 éléments pour être exact, dont seuls 8 sont obligatoires) permettant de répondre aux consignes édictées par ISAD(G). Si le nombre d'éléments descriptifs y est bien plus important, « c'est parce qu'il a été parfois nécessaire de répartir les informations entre plusieurs éléments pour leur manipulation informatique »³⁴, ce qui sous-entend que les normes archivistiques ont dû faire l'objet d'une adaptation pour permettre leur compatibilité avec l'environnement numérique.³⁵ La mise en œuvre de ce schéma de métadonnées répond à la volonté de mettre en ligne les instruments de recherche électroniques des archives, qui étaient auparavant au format papier. L'EAD (Encoded Archival Description) hérite de plusieurs caractéristiques techniques du langage XML, notamment en termes de structuration, permettant ainsi « d'utiliser et de manipuler le contenu des documents produits comme des données, d'échanger des données standardisées, de constituer de vastes réservoirs d'informations (portails archivistiques), de faciliter la recherche du public grâce à cette interopérabilité. »³⁶

Cependant, si la compatibilité de l'EAD avec les pratiques de description antérieures ne fait pas de doute, le recours à un tel schéma n'est pas sans inconvénient, au regard des besoins de diffusion des données issues de l'archivage et de l'interopérabilité avec d'autres schémas de métadonnées. En effet, l'EAD n'a par exemple pas été intégré au standard d'échange des données de l'archivage (SEDA) utilisé pour l'archivage électronique, comme le rappelle Claire Sibille de Grimoüard, sous-directrice du Service interministériel des Archives de France. En effet, il s'agit d'un outil développé sur mesure pour les besoins des archivistes, qui a été « conçu comme une transposition des habitudes issues du papier, enfermant ainsi les instruments de recherche dans une seule vision documentaire. »³⁷ La description en EAD doit-elle donc demeurer au cœur des pratiques ? De telles interrogations ne sont, là non plus, pas propres au monde des archives puisque les bibliothèques ont elles aussi été confrontées à la rigidité de leur

34 GRIMOÜARD. art. cit. p. 15

35 Cf. schéma représentant la structure de la DTD EAD en annexe 2.

36 GRIMOÜARD, Claire Sibille-De. Avantages et inconvénients d'un encodage en EAD. *La Gazette des archives*, 2010, Vol. 220, n° 4, pp. 179-194. DOI 10.3406/gazar.2010.4757.

37 Ibid.

langage de catalogage, avec le MARC³⁸ par exemple. Étienne Cavalié, responsable de l'analyse et du traitement des données à la BnF, dans un billet de blog de 2013³⁹, dresse le parallèle entre ces deux situations en rappelant que l'EAD peut encore jouer un rôle dans cette nouvelle ère de la description documentaire et archivistique résolument bouleversée par le web, mais en tant que format d'échange et non plus en tant qu'ensemble d'éléments à saisir manuellement. Dans les faits, c'est déjà ce qui se fait aujourd'hui dans les services d'archives, au sein desquels les archivistes ne sont parfois même pas conscients qu'ils sont en train d'encoder en EAD, puisqu'ils passent par l'intermédiaire de l'interface du logiciel métier et de ses masques de saisie.⁴⁰

En revanche, le schéma EAC-CPF, qui est à ISAAR(CPF) ce que EAD est à ISAD(G), a réellement été développé en même temps que la norme sur laquelle il repose. Tous les éléments d'ISAAR-CPF y sont donc intégrés et sa portée multidimensionnelle est alors préservée. Reposant également sur le langage XML, le schéma EAC « permet la production de documents pérennes, indépendamment de toute plate-forme logicielle »⁴¹ et échappe, dans une certaine mesure, aux restrictions liées aux instruments de recherche. Ceci étant, puisque ces derniers ont eu une influence considérable sur la manière dont a été pensée la description archivistique, dès ses prémices, il s'agit de nous intéresser à leur nature et aux mutations auxquelles ils sont eux aussi confrontés, d'autant plus que le recours aux schémas de métadonnées laisse entrevoir plusieurs conceptions de la description archivistique à l'ère du numérique.

I.3. Du document à la donnée, quelles implications pour les instruments de recherche ?

Document, ressource ou donnée : l'enjeu de la granularité

À la lumière des travaux du collectif Roger T. Pédaque concernant le document à l'ère du numérique, qui introduisent la notion de redocumentarisation⁴², c'est le régime de matérialité du document numérique et l'existence même de celui-ci en tant qu'unité de

38 MACHiNE-Readable Cataloging

39 CAVALIÉ, Étienne. L'EAD, c'est pour l'import-export. *Bibliothèques [reloaded]* [en ligne]. 7 octobre 2013. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://bibliotheques.wordpress.com/2013/10/07/lead-cest-pour-limport-export/>

40 Ibid.

41 GRIMOÛARD, Claire Sibille-de. Les normes internationales de description archivistique : origines, développements, perspectives. *La Gazette des archives*, 2012, Vol. 228, n° 4, pp. 77-90. DOI 10.3406/gazar.2012.4985.

traitement qui ont pu être remis en question. Traditionnellement, le document papier s'est vu caractérisé par un support et demeurerait alors un objet délimité physiquement, le contenu et la forme paraissant alors indissociables et donc traitées comme un tout par les professionnels de l'information. Ainsi, Jean-Michel Salaün qualifie ce changement de paradigme de la manière suivante :

« Le document est aujourd'hui un ensemble de pièces éparses réagencées à la volée, une tête d'un réseau concrétisée par les liens hypertextuels, et les traces laissées par les internautes sont des outils de gestion du trafic et des stratégies des firmes. Tous ses éléments constitutifs sont bousculés et la notion de document elle-même est ébranlée. »⁴³

Dès 2007, Bruno Bachimont emploie le terme d'éditorialisation pour qualifier la reconstitution de documents « sur mesure » à partir de données issues d'autres documents.⁴⁴ Le concept de redocumentarisation met en lumière l'importance des métadonnées pour enrichir les contenus, qui sont désormais « indispensables à la reconstruction à la volée de documents et toute la traçabilité de son cycle. »⁴⁵ Quoiqu'il en soit, cela implique une nouvelle manière de concevoir l'objet documentaire.

Marie-Anne Chabin, plus proche des préoccupations des archivistes, analyse ce tournant à travers le prisme de la double fonction du document, en faisant effectivement la différence entre le document « source » et le document « trace »⁴⁶. De ce fait, c'est pour elle davantage l'aspect formel du document qui s'en trouve profondément modifié, alors que la « démarche de production », et donc le contexte et la provenance, fondamentaux pour les archivistes, ne seraient que peu influencés par le numérique. Ceci étant, il demeure certain qu'avec l'essor du numérique, la notion de document est fragilisée et que c'est le concept de ressource, et désormais celui de

42 PÉDAUQUE, Roger T. *Le Document à la lumière du numérique : forme, texte, médium : comprendre le rôle du document numérique dans l'émergence d'une nouvelle modernité*. C & F Éditions, 2006. 187 p. ISBN 978-2-915825-11-4.

43 SALAÛN, Jean-Michel. Pourquoi le document importe. *INA Expert* [en ligne]. juin 2012. [Consulté le 26 août 2017]. Disponible à l'adresse : <https://www.ina-expert.com/e-dossier-de-l-audiovisuel-sciences-humaines-et-sociales-et-patrimoine-numerique/pourquoi-le-document-importe.html>

44 BACHIMONT, Bruno. Ingénierie des connaissances et des contenus : le numérique entre ontologies et documents. *Hermès Science Publ.*, 2007. ISBN 978-2-7462-1369-2.

45 SALAÛN, Jean-Michel. La redocumentarisation, un défi pour les sciences de l'information. *Études de communication. langages, information, médiations*, 1 octobre 2007, n° 30, pp. 13-23. DOI 10.4000/edc.428.

46 CHABIN, Marie-Anne. Document trace et document source. La technologie numérique change-t-elle la notion de document ? *Revue I3 – Information Interaction Intelligence* [en ligne], 5 juillet 2004, Vol. 4, n° 1. [Consulté le 21 avril 2017]. Disponible à l'adresse : https://archivesic.ccsd.cnrs.fr/sic_00001020/document

données, qui sont privilégiés. Pour les archivistes, il ne s'agit pas uniquement d'une redéfinition de la granularité de l'information à traiter, mais d'un enjeu en termes de préservation du patrimoine :

« Le numérique, s'il est susceptible d'être manipulé, d'effectuer de puissants calculs, d'être transmis et dupliqué, n'a pas vocation à être conservé sur le long terme : dans l'environnement numérique, un document n'existe plus en tant que tel. L'enjeu essentiel est donc le suivant : le document et son contexte doivent pouvoir être représentés dans un nouvel environnement culturel, social et technologique, un peu comme une composition musicale à partir d'une partition. »⁴⁷

Françoise Banat-Berger, directrice des Archives nationales, résume alors ce changement de paradigme de la manière suivante : « Habités à manipuler des objets physiquement délimités, qu'on les appelle fonds d'archives, versements ou dossiers, nous sommes désormais confrontés à la réalité de la donnée. »⁴⁸

L'instrument de recherche, document ou ensemble de données ?

Dans le monde des archives, où la « consignation d'un écrit certifié était le cœur des missions »⁴⁹, l'instrument de recherche constitue probablement l'indicateur le plus pertinent pour prendre la mesure de ces mutations. En tant que principal vecteur de matérialisation de la description archivistique, l'instrument de recherche constitue à la fois un outil incontournable pour l'archiviste et une ressource faisant encore office, sous des formes variées, d'interface permettant aux publics d'accéder au contenu des archives. L'instrument de recherche, que l'on trouve sous différentes formes en fonction des unités d'archives décrites, de l'état général des fonds au répertoire en passant par l'inventaire, voit lui aussi son existence en tant que document remise en question par les technologies du numérique et les évolutions en matière de description archivistique. Ainsi, il s'agit d'un point de départ intéressant pour analyser ces questions puisqu'il constitue, à la manière des métadonnées (qui sont, comme nous l'évoquons, « des données sur des données ») et des notices bibliographiques, « un document sur des documents. »⁵⁰ Il s'agit donc de déterminer si l'instrument de

47 BANAT-BERGER, Françoise. De l'écrit à internet : comment archive-t-on l'immatériel ? *Pouvoirs*, 20 avril 2015, n° 153, pp. 109-124.

48 BANAT-BERGER, Françoise et NOUGARET, Christine. Faut-il garder le terme archives ? Des « archives » aux « données ». *La Gazette des archives*. 2014. Vol. 233, n° 1, pp. 7-18. DOI 10.3406/gazar.2014.5121.

49 Ibid.

50 COLOMBANI, Louis. Instrument de recherche, données et document. *Anaphore* [en ligne]. 21 janvier 2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/instrument-de-recherche-donnees-et-document/>

recherche peut encore aujourd'hui réellement être considéré comme un document ou s'il doit être davantage perçu, en raison des évolutions technologiques que nous avons pu mettre en avant, comme un ensemble de données.

Dans un premier temps, la transition numérique et l'exposition de ces instruments de recherche s'est effectuée de manière plutôt basique, à travers la mise en ligne de documents PDF sur les sites des archives, ce qui n'a pas été sans poser problème.⁵¹ Il s'agit d'une solution qui a pu paraître viable dans la mesure où l'instrument de recherche a longtemps été conçu comme un document fini et publié sous forme de « livre » manipulable de manière isolée, mais qui ne correspond plus aux besoins et aux habitudes des internautes. En effet, comme nous avons pu le voir, les évolutions des outils numériques font que cette matérialité a perdu de son sens et que le potentiel, en matière de recherche et de réutilisation, de la mise en ligne des instruments de recherche serait alors perdu.⁵² Par ailleurs, si la nature documentaire de l'instrument de recherche imprimé ne fait pas vraiment de doute, la mise en œuvre des schémas de données que nous avons pu évoquer précédemment, tels qu'EAD et EAC-CPF, lors de leur conversion rétrospective ou de leur conception à partir d'un logiciel métier, peut laisser supposer que c'est le contenu d'un instrument de recherche, sous la forme de ces métadonnées descriptives, qui constitue réellement l'unité d'information à traiter et à exposer sur les portails et interfaces de consultation. L'instrument de recherche aurait alors tout intérêt à ne plus être « calqué sur le modèle imprimé » mais plutôt sur un modèle « ouvert, capable d'opérer en réseau et de dialoguer avec d'autres modèles de données. »⁵³ Céline Guyon pose alors la question suivante :

« Avec le numérique, les moteurs de recherche ne sont-ils pas en passe de supplanter les instruments de recherche ? Ou dit autrement : aura-t-on encore besoin d'instruments de recherche dans leur forme actuelle pour accéder aux archives numériques ? De clef d'accès aux fonds d'archives, les instruments de recherche ne pourraient-ils pas devenir les instruments de restitution des réponses à une

51 SIBILLE, Claire. Les instruments de recherche électroniques. *La Gazette des archives*, 2006, Vol. 204, n° 4, pp. 131-150. DOI 10.3406/gazar.2006.3831. « Non retravaillés dans une perspective Web, les documents PDF ne peuvent pas être complétés et modifiés, leur forme est figée, aucun lien ne peut être établi entre le contenu de divers instruments de recherche, il n'y a pas de recherche possible sur des descripteurs. En outre, le temps de chargement est très long pour l'internaute, d'où parfois la nécessité de découper les instruments de recherche en chapitres. »

52 Ibid.

53 ANGJELI, Anila. Bibliothèques, archives, identifiants : les enjeux de la normalisation. *Hermès, La Revue*, 25 novembre 2013, n° 66, pp. 167-172.

recherche. »⁵⁴

Louis Colombani, gérant de la société Anaphore, qui conçoit des logiciels métier et des outils à destination des services d'archives, vient nuancer cette observation en précisant que la complexité et la richesse des descriptions archivistiques, qui dépendent d'une structuration à plusieurs niveaux et d'une logique d'héritage, nécessitent une structuration qui dépend de leur destination, et la « restitution au seul format XML EAD » ne serait alors pas une « solution optimale ».⁵⁵ Il convient en effet de rappeler qu'avec les évolutions du web, et notamment la volonté d'aboutir à un web des données, concept que nous approfondirons ultérieurement, ces descriptions ne seront plus uniquement conçues pour être lues par des humains mais également pour être interprétées par des machines. D'après lui, il faudrait dissocier ces deux visions et penser les instruments de recherche et les moteurs de recherche comme des outils complémentaires.⁵⁶ Pour autant, l'instrument de recherche répond clairement à une culture de l'imprimé qui repose sur des « codes de représentation [...] intégrés dans un patrimoine commun, une culture commune »⁵⁷, ce qui ne signifie pas pour autant que ces codes soient facilement assimilés par les usagers auxquels les instruments de recherche sont destinés. En effet, dans leur version papier, les instruments de recherche ne sont pas toujours aisément pris en main par les publics et il s'agissait alors pour l'archiviste d'assurer en salle de lecture le rôle de médiateur.⁵⁸ Or, avec le numérique et l'exposition des contenus sur les sites web des services d'archives, c'est un contexte de désintermédiation qui s'est développé et il s'agit alors d'offrir la possibilité aux usagers de s'approprier les ressources et les données mises à disposition, de s'assurer que ces derniers soient en mesure de les comprendre et de les réutiliser. En effet, puisque les instruments de recherche de tous types ont été conçus dans le respect des pratiques de description et impliquent des concepts techniques qui découlent de pratiques métier, leur compréhension par les usagers peut être difficile, puisque « la plupart des internautes sont peu familiarisés avec la terminologie

54 GUYON. art. cit. p. 13

55 COLOMBANI. art. cit. p. 21

56 Ibid.

57 COLOMBANI, Louis. La description archivistique à l'ère du numérique – Part 1. *Anaphore* [en ligne]. 3 juillet 2014. [Consulté le 17 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/la-description-archivistique-a-lere-du-numerique-part-1/>

58 BROCHIER, André. Instruments de recherche et Internet : référentiels et métadonnées. *La Gazette des archives*, 2010, Vol. 220, n° 4, pp. 113-125. DOI 10.3406/gazar.2010.4745.

archivistique et la structure des instruments de recherche [...] et que les modalités de navigation au sein de la hiérarchie des niveaux de description ne sont pas toujours bien comprises des utilisateurs. »⁵⁹ Ainsi, pour Claire Sibille de Grimoüard, « une plus grande attention portée aux utilisateurs, avec des facilités offertes pour la consultation et l'interrogation des instruments de recherche, devra donc être au cœur des préoccupations des prochaines expériences de mise en œuvre des normes et formats d'encodage. »⁶⁰ Ce sont effectivement les questions de la diffusion et de l'accès qui doivent alors être étudiées au prisme des évolutions du web et des pratiques des internautes, étant donné que « les archives, comme les bibliothèques, ont depuis longtemps perdu ce qui en faisait des lieux uniques parce qu'incontournables pour accéder aux sources primaires et donc au savoir »⁶¹ C'est alors l'ensemble de la chaîne traditionnelle de l'archivage, de la collecte à la communication, qui doit être reconfigurée, pour reprendre l'expression de Céline Guyon, afin d'assurer une visibilité des documents d'archives sur le web.

Pour autant, si le numérique bouleverse effectivement la conception de l'instrument de recherche, il constitue une opportunité, à travers l'utilisation de schémas de métadonnées et l'émancipation par rapport à un carcan documentaire trop rigide pour être interopérable, pour les archivistes mais également pour les publics. Si l'instrument de recherche constitue un exemple pertinent pour aborder cette transition, il s'agira pour nous de transposer cette problématique à l'échelle de l'ensemble des pratiques des archivistes et au contexte de la mise en place d'un système d'archivage électronique, en analysant notamment en quoi cette transition vers la donnée influe sur le métier d'archiviste, dont les compétences tendent alors à se rapprocher de celles habituellement associées à d'autres professionnels de l'information.

59 GRIMOÛARD. art. cit. p. 19

60 Ibid.

61 GUYON. art. cit. p. 13

II. De l'élaboration de modèles conceptuels à la mise en place de référentiels : les enjeux d'une gouvernance de l'information et des données

II.1. La gestion de l'information à l'ère du numérique, une fonction transversale

La gouvernance, un concept défini par son caractère systémique

Le concept de gouvernance, dans son acception la plus courante, renvoie à une forme de management transversal issue des théories des sciences administratives et de la gestion, d'influence principalement anglo-saxonne.⁶² Souvent difficile à définir, il suppose une efficacité en matière de prise de décision et une implication des organes décisionnels et des managers dans les phases opérationnelles des projets.⁶³

L'expression de gouvernance de l'information, apparue quant à elle au tournant des années 2010, s'inscrit dans une logique déjà bien installée dans le monde de la gestion de l'information et du document, en faisant référence à un échelon supérieur à ce qui a longtemps été nommé le management de l'information.⁶⁴ Cette nouvelle conception de la gestion de l'information à l'échelle d'une organisation repose sur une vision systémique et une approche globale de l'information au sein d'une organisation.⁶⁵ Arnaud Jules, directeur de la gestion et de la conservation de l'information de la société Orange, et Loïc Lebigre, consultant à Centre Inffo, rappellent que si elle est « restée souvent orale et implicite, l'information fait l'objet de processus de mise en écrits. »⁶⁶ De ce fait, l'information constituerait alors « une ressource, tant cognitive (à la fois agent et résultat de la réflexivité des acteurs puisqu'elle modifie leurs représentations antérieures) que décisionnelle et stratégique. »⁶⁷

62 JULES, Arnaud et LEBIGRE, Loïc. Information : le temps de la gouvernance. *Documentaliste-Sciences de l'Information*, 15 avril 2013, Vol. 50, n° 1, pp. 24-25

63 Ibid. « La gouvernance signifie le paramétrage des politiques, des règles, de l'organisation, des processus et des contrôles de l'entreprise pour maintenir l'entreprise en conformité avec toutes ces exigences sous le régime d'un cadre de gouvernance d'entreprise. »

64 GUYOT, Brigitte, FUZEAU, Pierre, COTTIN, Michel et RIETSCH, Jean-Marc. Repères conceptuels. *Documentaliste-Sciences de l'Information*, 15 avril 2013, Vol. 50, n° 1, pp. 26-37

65 D'après la définition du groupe SERDA : « la stratégie et l'organisation définies afin de gérer, sécuriser, partager, conserver, valoriser et maîtriser l'information et la connaissance dans l'organisation et son environnement externe afin d'améliorer l'efficacité d'une organisation.

66 JULES, LEBIGRE. art. cit. p. 25

67 Ibid.

Si la notion semble bel et bien être originaire du milieu de la gestion documentaire, elle dépasse désormais ce seul domaine et concerne l'ensemble des professionnels qui travaillent au quotidien avec la ressource informationnelle⁶⁸, tout en intégrant le rôle des décideurs. Les enjeux sont clairs, puisqu'il s'agit d'accroître l'efficacité des décisions et, dans une approche héritée du records management, de réduire les risques ainsi que les coûts pour l'organisation. Ces enjeux ont pris une toute nouvelle dimension avec l'essor du numérique et un certain nombre d'opportunités a pu émerger, puisque l'un des objectifs d'une gouvernance unifiée est la sortie d'une logique de silos d'information et le décloisonnement par la mise en œuvre d'un niveau d'interactions plus important. La gouvernance de l'information dépasse donc les compétences d'un seul corps professionnel et a pour objet les interactions entre les acteurs concernés par la manipulation de l'information en tant que ressource stratégique. Ainsi, pour développer une gouvernance de l'information efficace, une organisation devrait s'extraire de la logique selon laquelle l'information est l'apanage d'une seule catégorie de professionnels.⁶⁹ En vertu de cette nouvelle vision stratégique de l'information, les missions documentaires des organisations font l'objet de mutations qui vont dans le sens d'une optimisation des processus métiers pour répondre à des besoins organisationnels. Dans les faits, et afin de faire le parallèle avec notre objet d'étude, il s'agira pour nous d'analyser, à la lumière des rapprochements entre ces différents corps de métier que l'environnement numérique favorise en termes de mutualisation des démarches et des procédures, l'impact de cette vision globale sur la pratique de l'archivage et plus précisément l'influence des sphères de l'informatique et du management sur le champ réglementaire et les processus métier propres à l'archivage.

Les apports des compétences de l'archiviste à cette vision globale

Si la GI est effectivement le fruit d'une gestion stratégique des processus informationnels et repose sur une instance décisionnelle forte, elle est également caractérisée, comme nous l'évoquions, par une transversalité accrue. En effet, Jürg Hagmann et Jean-Daniel Zeller, consultants en records management et archivage

68 HAGMANN, Jürg et ZELLER, Jean-Daniel. Gouvernance de l'information : véritable innovation dans la gestion de l'information ? *Les Cahiers du numérique*, 17 juin 2015, Vol. 11, n° 2, pp. 15-36.

69 Logique qui a pourtant la peau dure dans de nombreuses structures où les services informatiques se sont accaparé un rôle central. cf. Ibid.

électronique en Suisse, insistent sur cet aspect en proposant la définition suivante : « la GI est l'art de l'interaction de confiance entre les principales parties prenantes d'un programme de GI (informatique, métiers, juridique et conformité, RIM, sécurité et confidentialité). » Puisqu'il s'agit alors de favoriser les interactions entre ces différents acteurs, « le plus grand défi sur le chemin de la gouvernance unifiée est le fait qu'aucune fonction ou service seul n'est en mesure d'atteindre les objectifs et les avantages perçus »⁷⁰, ce qui implique une redistribution des cartes, sur laquelle ces mêmes auteurs insistent en signalant que la GI « dépend d'une culture de la gestion du changement efficace et d'une extrême collaboration interfonctionnelle. »⁷¹

D'après la dernière édition du rapport du groupe SERDA sur la gouvernance de l'information numérique dans les organisations⁷², la GI est devenue une réalité pour 75 % des organisations et semble prendre de l'ampleur pour les entreprises comme pour les administrations. Pour autant, ce même rapport indique que les métiers de l'information-documentation, en tant que pilotes de projets de GI, sont en net recul et que les services d'archives ne sont que rarement placés à la tête de ces politiques de gouvernance informationnelle. En effet, ce sont davantage les directions des systèmes d'information et les directions générales qui s'emparent de ces aspects, ou bien des « instances dédiées » ou encore des « binômes de coordinateurs »⁷³, ce qui pose la question de la façon dont les archivistes sont intégrés à ces dispositifs.

Au regard des impératifs de l'archivage électronique et des besoins en termes d'authenticité et de pérennité des documents et données à conserver, les compétences traditionnelles des archivistes, en matière d'élaboration de tableaux de bord et de plans de classement reposant sur des typologies documentaires et répondant à des besoins d'archivage spécifiques repérés en amont, peuvent se voir transposer à cette nouvelle dimension systémique. La gouvernance de l'information constitue alors un contexte « favorable aux archivistes qui sont à même d'apporter leurs réalisations et réflexions en termes d'authenticité des archives, de référentiels et contextualisation de

70 HAGMANN, ZELLER. art. cit. p. 26

71 Ibid.

72 SERDA, 2017. La Gouvernance de l'information numérique dans les organisations. [en ligne]. mai 2017. [Consulté le 12 août 2017]. Disponible à l'adresse : <http://www.serda.com/conseil/transition-num%C3%A9rique/gouvernance-de-l%E2%80%99information-%E2%80%93-un-outillage-de-plus-en-plus>

73 Ibid.

l'information (métadonnées descriptives, thesaurus, classement et évaluation des archives). »⁷⁴ Par ailleurs, le développement de systèmes de gestion électronique de documents, comme le signale Françoise Banat-Berger, a déjà permis aux archivistes de contribuer très largement au processus de dématérialisation. En effet, elle rappelle que « les archivistes sont les acteurs de cette dématérialisation puisque, sans eux, les choix opérés auraient été très différents (périmètre de la numérisation, choix de pièces, réflexion sur la valeur juridique des pièces) »⁷⁵, tout en émettant l'hypothèse qu'en leur absence, de nombreuses questions relatives au cycle de vie de l'information et à la stratégie d'archivage auraient tout simplement été éludées. Jean-Marc Rietsch appuie sur ce point en affirmant que « tant la dématique que l'archivage électronique sont directement concernés par ce véritable défi lié à la gouvernance de l'information, comme partie intégrante du système d'information »⁷⁶, ce qui souligne la nécessité pour les archivistes de s'emparer de ces questions et, dans une optique de mutualisation des compétences, de faire profiter aux autres acteurs d'une gouvernance de l'information de leurs acquis en matière de conservation et de traitement des documents.

La collaboration avec les directions informatiques, l'occasion d'une réévaluation du positionnement de l'archiviste

La nature hybride de l'archivage, avec des documents papier et des ressources électroniques, qui doivent bénéficier de la même attention, implique par ailleurs une collaboration avec les informaticiens que Françoise Banat-Berger qualifie d'essentielle. Celle-ci ajoute par ailleurs que « l'archiviste doit ainsi être en mesure d'intervenir dès la conception des applications métier ou lors de leur évolution de manière à définir et mettre en œuvre le cycle de vie des informations gérées ainsi qu'une stratégie d'archivage. »⁷⁷ Dès lors, ce sont également de nouvelles compétences mais également une certaine forme d'acculturation aux aspects informatiques de l'archivage qui pourraient caractériser, dans une certaine mesure, les évolutions à venir du métier d'archiviste.

74 BANAT-BERGER, Françoise. De l'écrit à internet : comment archive-t-on l'immatériel ? *Pouvoirs*, 20 avril 2015, n° 153, pp. 109-124.

75 Ibid.

76 GUYOT et al. art. cit. p.25

77 BANAT-BERGER. art. cit. p. 28

Ces évolutions laissent alors supposer que c'est la posture de l'archiviste, en raison de cette nouvelle articulation mais également parce que les impératifs de la dématérialisation et de la « e-administration » imposent un accroissement exponentiel en termes de volumétrie documentaire et une prise en compte accrue de l'aspect stratégique de la donnée, qui pourrait être modifiée en profondeur. Dans les faits, ce contexte nécessite une remise en question du positionnement de l'archiviste, qui se voit contraint « d'adapter ses modes d'intervention face aux nouveaux modes de production documentaire et, d'autre part, participer et tirer les conséquences de la dématérialisation du processus métier. »⁷⁸ Ainsi, Céline Guyon rappelle qu'il est « dangereux de croire que la production même des documents ne nécessite pas les compétences d'un professionnel qui puisse l'orienter, la contrôler, créer les référentiels les plus adaptés au contexte » et qu'une intervention « a posteriori » n'est désormais plus suffisante. Il s'agit alors d'un tournant inédit dans la pratique de l'archivage, puisque « l'archiviste participe au mode de constitution de l'information et intervient avant même que l'information ne soit créée. »⁷⁹ Ce constat rejoint notre interrogation initiale sur la validité de la théorie des trois âges et de la conception, propre à l'archivistique en France, du cycle de vie de l'information.

II.2. L'interopérabilité : condition première de la gouvernance des données

Les outils de la gouvernance : de la politique d'archivage au référentiel de données

Dans une démarche de GI, il appartient aux décideurs, aux responsables, de proposer les moyens de faire émerger une réelle politique documentaire, ou politique d'archivage, et d'en assurer le maintien, à l'échelle d'une organisation. Cette politique est le véritable point de départ de la gouvernance de l'information, délimitant « des objectifs précis pour lesquels une stratégie est déployée. »⁸⁰ Jean-Marc Rietsch, expert en dématique et dirigeant de la société FISAM, en souligne l'importance de la façon

78 BANAT-BERGER, Françoise. « Un métier à part entière, l'archiviste un généraliste de l'information » : qu'en est-il en 2012 dans le nouvel environnement numérique ? *La Gazette des archives*, 2012, Vol. 226, n° 2, pp. 117-126. DOI 10.3406/gazar.2012.4901.

79 Ibid.

80 JULES, Arnaud, MOURAIN, Jean, BLANGER, Jean-Pierre, MASSE, Claudine et CHABIN, Marie-Anne. Pratiques. *Documentaliste-Sciences de l'Information*, 15 avril 2013, Vol. 50, n° 1, pp. 50-59.

suivante :

« La politique documentaire et d'archivage de toute organisation pose en fait les bases de sa politique en matière de gestion de l'information et des engagements qu'elle prend tant en interne que vis-à-vis de l'extérieur. Elle représente la garantie d'une démarche cohérente et contrôlée face à l'ensemble des exigences qui s'imposent à l'organisation en regard des besoins de sécurité et de ses besoins métier. »⁸¹

En outre, le champ normatif de la gouvernance de l'information ne repose plus uniquement sur des normes techniques mais également sur des normes dites de « haut niveau », telles que les normes relatives au management de la qualité, (comme la norme ISO 9001).⁸² Dans la pratique, ce sont les processus métiers⁸³ qui doivent faire l'objet d'une harmonisation. La mise en œuvre d'une gouvernance de l'information passe alors en premier lieu par une « cartographie des processus [...] propres à la gestion de l'information et des documents ayant été au préalable listés dans la politique documentaire. »⁸⁴ D'après Jean-Pascal Perrein, fondateur du projet GouvInfo IAI, le préalable absolument nécessaire à la mise en œuvre d'une GI est « la connaissance des référentiels relatifs aux informations de l'organisation : les référentiels de données, de conservation, d'informations confidentielles, d'informations personnelles, etc. »⁸⁵ C'est d'après lui à un objectif de « maîtrise de l'information »⁸⁶ sur l'ensemble du cycle de vie que doit répondre l'instance de gouvernance de l'information. En effet, les différents référentiels propres à une organisation, et plus particulièrement, en raison de notre objet d'étude, les référentiels de données et référentiels de patrimoine, constituent un élément crucial de l'architecture informationnelle d'une organisation.

Pour ne pas rester dans le domaine du vœu pieux, et puisqu'elle ne constitue pas une technique en elle-même, la GI doit être appuyée par un ensemble d'outils de natures

81 GUYOT et. al. art cit. p. 25

82 JULES, LEBIGRE. art. cit. p. 25

83 Toujours d'après JULES et LEBIGRE : « suite d'opérations normalisées effectuées par tout ou partie des collaborateurs d'une fonction métier de l'organisation pour effectuer une activité donnée »

84 JULES, et. al. art. cit. p. 25

85 PERREIN, Jean-Pierre. Politiques et référentiels de la gouvernance de l'information. *GouvInfo IAI* [en ligne]. 29 mai 2012. [Consulté le 10 août 2017]. Disponible à l'adresse : <http://www.gouvinfo.org/IAI/politiques-et-referentiels-de-la-gouvernance-de-linformation/>

86 Cf. Ibid. « Tout comme les référentiels, elle est garante de la cohérence d'ensemble de ces politiques, ainsi que de leur lisibilité et de leur exploitabilité. Un trop grand nombre de nos politiques sont aujourd'hui illisibles et/ou inexploitables pour les utilisateurs à qui elles s'adressent. Les rendre accessibles d'une part, et cohérentes entre elles d'autre part, favorise grandement une meilleure maîtrise de l'information. »

différentes⁸⁷: des outils de pilotage aux outils de soutien en passant par les outils de gestion et autres modèles de référence. Parmi ceux-ci, ce sont plus particulièrement les outils de gestion, à savoir les « outils documentaires [...] définis dans les normes de gestion des documents d'activité »⁸⁸ qui vont nous intéresser. Ils regroupent en effet des procédures et techniques qui sont le plus souvent le fruit du travail des professionnels de l'information, et notamment des archivistes.⁸⁹ Parmi ces outils de gestion, on trouve notamment le plan de classement, les chartes de nommage et d'indexation, ainsi que les règles destinées à assurer la gestion de l'identification pérenne des documents. L'identification pérenne de l'origine et de la propriété des documents ainsi que la détermination des durées de conservation appropriées pour les documents figurent également parmi ces compétences documentaires indispensables.

Le recours à des systèmes d'organisation des connaissances et autres référentiels terminologiques rejoint également les besoins de transversalité et d'interopérabilité qui caractérisent la GI. En effet, Pierre Fuzeau, président de la CN46 à l'AFNOR, rappelle que « la gouvernance technique de l'info-connaissance nécessite des référentiels que l'on a trop hâtivement et rapidement mis au rebut » et précise dès lors que « des outils sont à rechercher vers les normes d'ontologies conçues par le W3C comme le SKOS (Simple knowledge organisation system). »⁹⁰ Parmi ces systèmes d'organisation des connaissances, les vocabulaires contrôlés (thésaurus, taxonomies, liste d'autorités, etc.), qui font l'objet d'une application de longue date par les professionnels de l'information, archivistes compris, sont aux premières loges de ces impératifs d'harmonisation de l'information au sein d'une structure, dans la mesure où ils sont les garants d'une rationalisation des usages, fournissant à une communauté d'utilisateurs un vocabulaire commun. Dès lors, il n'est pas étonnant de voir le groupe SERDA proposer, parmi ses formations, un séminaire sur la gestion de vocabulaires contrôlés. La page dédiée à celui-ci rappelle justement que ces outils « prennent une importance croissante dans les systèmes d'information [et] jouent un rôle critique pour l'interopérabilité et la connexion des jeux de données, leur réutilisation et leur

87 JULES, Arnaud. Une politique de gestion des documents d'activité pour une gouvernance documentaire stratégique. *La Gazette des archives*, 2012, Vol. 228, n° 4, pp.153-171. DOI 10.3406/gazar.2012.4991.

88 Ibid.

89 Ibid.

90 GUYOT et al. art. cit. p.25

archivage », et pose leur gouvernance comme un prérequis d'un « dialogue permanent avec les métiers, la DSI et les utilisateurs internes ou externes. »⁹¹

Modèles conceptuels et standards d'échanges : l'archivage électronique dans un contexte de gouvernance

Dans la mesure où la politique d'archivage peut être considérée comme l'un des fondements d'une gouvernance de l'information et puisque les systèmes de GED et les SAE en constituent le socle technique, il convient de nous intéresser aux outils qui garantissent la stabilité de ces systèmes. En effet, la multiplication des projets d'archivage électronique a mené à un effort de normalisation, dès le milieu des années 2000, pour répondre aux impératifs de pérennisation des contenus électroniques et pour assurer la valeur probante et l'intégrité de ces derniers. En raison de leur nombre et de leur diversité, et puisqu'elles ne sont qu'une facette de notre sujet, nous ne nous attacherons pas à les présenter dans toute leur exhaustivité.⁹² Il convient néanmoins, afin d'être en mesure d'évaluer leur impact sur les pratiques, d'évoquer celles qui s'inscrivent dans ce besoin de transversalité des processus et d'interopérabilité qui caractérise la GI. Comme le souligne Stéphanie Roussel, conservateur du patrimoine au SIAF, les normes de l'archivage électronique constituent un socle indispensable, dans la mesure où elles « établissent un fondement stable et relativement pérenne dans un environnement technologique qui ne cesse de se transformer, [...] permettent aux projets d'archivage électronique de s'adosser à des socles partagés et reconnus. »⁹³

Si ces normes et ces outils fournissent un cadre de travail garantissant une gestion documentaire robuste, c'est le modèle OAIS (Open Archival Information System), résultat du travail du CCSDS (Consultative Committee for Space Data Systems) décrit dans la norme ISO 14721, qui « spécifie les composantes d'un système apte à assurer la préservation et l'accès à long terme aux documents. »⁹⁴ Il s'agit d'un modèle interopérable avec les normes précédemment citées, qui présente l'architecture

91 SERDA. Gérer et publier des référentiels et des thésaurus. Formation SERDA [en ligne]. 3 mars 2015. [Consulté le 1 septembre 2017]. Disponible à l'adresse : http://www.formation-serda.com/formations-documentation/gerer_publier_referentiels_thesaurus_0

92 Cf. schéma du paysage normatif de l'archivage électronique en annexe 3.

93 ROUSSEL, Stéphanie. Le champ normatif de l'archivage électronique. *La Gazette des archives*, 2012, Vol. 228, n° 4, pp. 59-76. DOI 10.3406/gazar.2012.4984.

94 DHÉRENT, Catherine, GRIMOÛARD, Claire Sibille De, OTT, Florence et BROULAND, Myriel. Comprendre les enjeux et les fondamentaux. *Documentaliste-Sciences de l'Information*. 5 juin 2009. Vol. 46, n° 2, pp. 32-42.

logique et les interactions sur lesquelles va pouvoir reposer le système d'archivage. En effet, ce sont les différentes opérations qui vont avoir lieu entre les quatre catégories d'acteurs, à savoir l'opérateur d'archivage, les producteurs, le management et les utilisateurs, qui y sont spécifiées. Afin de permettre la gestion d'un flux de données, l'information y est découpée en paquets, qui dépendent de leur localisation dans l'archive et qui réunissent fichiers à préserver et métadonnées nécessaires à l'archivage.⁹⁵ La norme précise par ailleurs l'ensemble des types d'informations qui doivent être préservées avec les objets archivés, à savoir des métadonnées de description, de représentation, de pérennisation et d'empaquetage. C'est sa nature ouverte qui fait que c'est sur l'OAIS « que s'appuient les politiques d'archivage [...] et la plupart des plates-formes d'archivage électronique conçues ces dernières années. »⁹⁶ Dans la mesure où le modèle OAIS intègre les processus décisionnels, essentiels pour le bon fonctionnement du système, qui sont la responsabilité de l'instance « management », il s'inscrit dans cette logique de gouvernance de l'information. En effet, « pour le management, le système d'archivage s'insère dans un plan stratégique global [...] C'est évidemment au management qu'il revient de soutenir le dispositif, politiquement, financièrement, et sur le très long terme. »⁹⁷ Il s'agira alors pour nous, puisque ce modèle a avant tout pour finalité de représenter les interactions entre les différents acteurs de l'archivage, de nous intéresser à la manière dont les archivistes y conçoivent leur positionnement et de déterminer si ces outils font sens pour les services d'archives qui dépendent encore largement des pratiques papier.

La définition d'un modèle conceptuel n'est cependant pas suffisante au bon déroulement des interactions entre les différents acteurs qui y sont définis, puisque pour être en mesure d'échanger des données, ceux-ci doivent partager, en plus des processus et de la terminologie, un ensemble de standards d'échange. C'est là qu'intervient désormais le Standard d'échange des données pour l'archivage (SEDA), qui « modélise les différentes transactions qui peuvent avoir lieu entre des acteurs

95 Les SIP (Submission Information Package), c'est-à-dire les paquets versés par le producteur, les AIP (Archival Information Package), à savoir les données archivées et enfin les DIP (Diffusion Information Package), qui correspondent aux données diffusées. Cf. schéma en annexe 4.

96 DHÉRENT et. al. art. cit. p.32

97 CENTRE INFORMATIQUE NATIONAL DE L'ENSEIGNEMENT SUPÉRIEUR. Le modèle de référence : l'OAIS. CINES [en ligne]. [Consulté le 17 avril 2017]. Disponible à l'adresse : <https://www.cines.fr/archivage/un-concept-des-problematiques/le-modele-de-reference-loais/>

dans le cadre de l'archivage de données »⁹⁸. La nature générique du SEDA assure sa compatibilité avec l'ensemble des types de documents de données, qu'elles soient électroniques ou papier, et favorise l'interopérabilité entre les systèmes d'archivage des services publics, ce qui permet une meilleure mutualisation des réalisations informatiques. Le recours à des standards de métadonnées interopérables et reposant sur le langage XML est dès lors un élément crucial de l'archivage électronique. Le standard PREMIS⁹⁹ fournit les lignes directrices pour assurer la pérennité des objets numériques et garantir leur compréhension sur le long terme. Ce dictionnaire de données identifie les entités impliquées dans la conservation des documents numériques¹⁰⁰, les relations entre ces entités et les métadonnées qui doivent être présentes. Il s'agit là aussi d'un outil interopérable avec les normes précédemment évoquées, et notamment avec le modèle OAIS, sur lequel il s'appuie fortement.

Dans l'idée, ces différents dispositifs, à savoir ces normes, ces modèles conceptuels et autres standards d'échanges ont pour vocation d'uniformiser la pratique afin de faciliter les interactions entre les acteurs de l'archivage. Dans la pratique, en revanche, ils pourraient représenter un enchevêtrement complexe, dont les enjeux seraient difficiles à percevoir par les archivistes, qui doivent désormais être en mesure de dialoguer avec d'autres catégories de professionnels, notamment les informaticiens, tout en prenant en compte les pratiques existantes pour les documents papier. Ainsi, malgré sa diversité, le cadre normatif de l'archivage électronique demeure « difficile à appréhender en raison de la diversité des points de vue et des périmètres couverts, mais également en raison des recouvrements partiels existant entre certaines normes. »¹⁰¹ En raison de ces difficultés, Stéphanie Roussel préconise un rapprochement entre les différentes entités dont les missions sont concernées par

98 Archives électroniques | Normes et standards | SEDA. In : *FranceArchives* [en ligne]. [Consulté le 19 août 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/26287460>. Pour ce faire, le SEDA repose sur un ensemble de scénarios, présentés sous la forme d'un diagramme UML, et précise les messages que les acteurs doivent s'échanger, par le biais de schémas XML, tout en définissant la manière de coder ces informations (c'est-à-dire sous la forme de métadonnées). Ces scénarios et ces messages concernent l'ensemble de la chaîne de l'archivage, de la demande de transfert à l'élimination ou à la communication. Le standard comprend aussi des balises concernant le sort final et la durée d'utilité administrative d'une archive.

99 Dictionnaire de données de métadonnées de pérennisation. Cf. <https://www.loc.gov/standards/premis/>

100 À savoir l'entité intellectuelle, l'objet, l'évènement, l'agent et les droits.

101 ROUSSEL. art. cit. p.31

l'archivage numérique, et signale alors « qu'une approche transverse serait probablement bénéfique pour tous. »¹⁰²

Cette deuxième partie nous aura permis de mettre en lumière les exigences croisées de la gouvernance de l'information et de l'archivage électronique, qui doivent reposer sur des pratiques transversales et des référentiels de données robustes pour assurer la pérennité et l'authenticité des documents et des données. Il s'agira pour nous d'évaluer si ces mutations sont réellement perçues par les professionnels des archives ou si le cloisonnement en termes de compétences est toujours un état de fait. La participation de l'archiviste à la mise au point de référentiels constituerait pour lui une chance de devenir un acteur d'une gouvernance des données. Ce faisant, il s'agirait d'assurer la qualité et la fiabilité des données, préalable nécessaire à leur éventuelle ouverture et, par extension, à leur exposition sur le web.

III. Optimiser la gouvernance dans un service d'archives : les potentialités des technologies du web sémantique

III.1. Les fondements du web sémantique : créer du lien et donner du sens aux données

Le web sémantique, une vision présente dès les prémices du web

C'est le constat d'une « absence de cadre d'interopérabilité pour échanger dans un espace de machines en réseau les documents et les données contenus dans les ordinateurs »¹⁰³ qui a constitué le point de départ du projet du web, tel qu'envisagé par T. Berners-Lee et qui est désormais qualifié de web « 1.0 », ou web des documents. Pour autant, dès la première conférence WWW en 1994 et la création du W3C la même année, Tim Berners-Lee insiste sur la nécessité de prévoir une dimension « sémantique » du web.¹⁰⁴ Projet éminemment ambitieux, il a effectivement pour objectif de permettre aux « machines », c'est-à-dire aux différents dispositifs (logiciels, algorithmes, etc.) qui vont pouvoir accéder et exploiter les données du web, de construire du sens, à partir des ressources mises en ligne afin de permettre une

102 Ibid.

103 POUPEAU, Gautier. Petite histoire du Web sémantique. *Les petites cases* [en ligne]. août 2011. [Consulté le 3 mai 2017]. Disponible à l'adresse : <http://www.lespetitescases.net/petite-histoire-du-web-semantique>

104 Ibid.

restitution qui coïnciderait avec les processus cognitifs et modèles de pensée humains. L'idée repose sur le fait qu'un internaute « ne parcourt pas le Web de façon aléatoire, il mobilise les modèles qu'il a du monde. »¹⁰⁵ Il s'agirait alors de comprendre ces modèles, d'en construire des représentations formelles permettant d'établir des liens entre les ressources, dans l'optique d'une amélioration des interactions avec le web. Cette vision de ce que doit être le web, avec des ressources et des données interconnectées qui ne dépendent plus du carcan documentaire et de pages hypertexte, constitue alors le fameux « web sémantique », tel que Tim Berners-Lee le baptise dans son article publié en 2001.¹⁰⁶ Le web sémantique, tel qu'il y est présenté, ne constitue alors pas une nouvelle version du web, mais plutôt une brique supplémentaire qui va reposer en grande partie sur la richesse de l'architecture déjà en place.

Si c'est une conception des évolutions technologiques qui existe depuis les débuts du web, il aura fallu attendre une dizaine d'années avant que les standards du W3C liés à ces questions et proposant des solutions techniques pour mettre en œuvre ce web de données ne fassent leur apparition. Dans l'objectif d'aboutir à des réalisations plus concrètes de cette vision initiale, le projet de Linked Open Data est lancé par le W3C et propose le cadre, via un certain nombre de recommandations et de règles¹⁰⁷, de l'émergence de « données liées et ouvertes » sur le web. Pour voir l'arrivée de premières interfaces qui utilisent pleinement ces technologies, en concordance avec les standards édictés par le W3C, il faudra attendre 2007, avec la création de Dbpedia¹⁰⁸, un gisement de données créé par deux universités allemandes et reposant sur des données structurées extraites de Wikipédia.

La mise en œuvre d'un projet qui s'inscrit dans cette vision du web exige une articulation entre les contenus ou ressources qui sont mis en ligne sur le web et les modèles et schémas de description communs qui permettront d'établir des liens entre

105 BACHIMONT, Bruno, GANDON, Fabien, POUPEAU, Gautier, VATANT, Bernard, TRONCY, Raphaël, POUYLLAU, Stéphane, MARTINEZ, Ruth, BATTISTI, Michèle et ZACKLAD, Manuel. Enjeux et technologies : des données au sens. *Documentaliste-Sciences de l'Information*. 31 janvier 2012. Vol. 48, n° 4, pp. 24-41.

106 BERNERS-LEE, Tim, HENDLER, James, LASSILA, Ora, et al. The semantic web. *Scientific american*, 2001, vol. 284, no 5, p. 28-37.

107 Cf. Le programme de déploiement en cinq étoiles du Linked Open Data : <http://5stardata.info/en/>

108 <http://wiki.dbpedia.org/>

des données d'horizons divers. Ainsi, « la transition du web sémantique vers le web de données se fera par qualification desdites données grâce à des schémas de description communs qui permettront de promouvoir des données réellement liées. »¹⁰⁹ La première couche technologique permettant cette qualification est l'URI (Unified Resource Identifier), un identifiant unique permettant, comme son nom l'indique, de nommer et d'indiquer toute ressource présente sur le web.¹¹⁰ Le fait d'identifier cette ressource est l'ingrédient essentiel de l'interconnexion des données et se trouve à la base du « layer cake »¹¹¹ des technologies du web sémantique. L'infléchissement, d'un web sémantique visant à décrire les ressources présentes sur le Web à l'aide de métadonnées, vers un web des données davantage tourné vers l'exposition massive de ces données, constitue le point de départ de la conception d'un ensemble d'outils et de recommandations, dont le modèle RDF est le socle.

Ainsi, comme le précise Bruno Menon, RDF (Resource Description Framework) « se voulait initialement un modèle général de métadonnées, permettant de décrire les caractéristiques des documents du Web. »¹¹² Ce modèle postule que décrire une ressource, c'est renseigner à son sujet une ou plusieurs caractéristiques, c'est-à-dire des propriétés munies de leurs valeurs. Pour ce faire, le RDF a recours à une logique de triplets d'information (sujet, prédicat, objet). Le modèle RDF et les triplets qui le composent bénéficient alors du mécanisme des URI afin de déclarer l'ensemble des relations qui peuvent exister entre les ressources identifiées par ces derniers. La nature ouverte et interopérable de ces technologies, et le fait qu'elles reposent sur l'existant robuste des webs documentaire et social, permet alors que « de telles déclarations peuvent provenir de n'importe quelle source sur le Web et être fusionnées avec d'autres, ouvrant ainsi des perspectives d'intégration de données à l'échelle mondiale. »¹¹³

109 CHARLET, Jean et KEMBELLEC, Gérald. Du web sémantique au web des données, quels enjeux professionnels ? *I2D – Information, données & documents*, Vol. 53, n° 2, pp. 54-55.

110 BACHIMONT et al., art. cit p. 37 Ces ressources peuvent alors « être absolument n'importe quoi, on peut donc traduire URI par « identifiant d'une chose » réelle, imaginaire, concrète, abstraite, individuelle ou générale : une personne, une organisation, un lieu, un événement, ou encore un concept dans un vocabulaire. »

111 Également appelé Semantic Web Stack. Cf. schéma en annexe 5.

112 MENON, Bruno. Comprendre les standards du web de données. *I2D – Information, données & documents*, février 2016, Vol. 53, n° 2, pp. 32-34.

Ontologies et vocabulaires : le socle de la sémantisation des données

Pour uniformiser et garantir l'interopérabilité de la description des ressources, les éléments qui composent celle-ci et les relations qui peuvent être établies entre ces différents éléments sont spécifiées à l'aide de « vocabulaires », qui ne doivent pas être confondus avec les vocabulaires contrôlés, puisqu'ils sont en réalité des ontologies permettant de « contraindre et documenter les moyens descriptifs utilisables »¹¹⁴. Les ontologies jouent un rôle déterminant dans la mise en place du web des données, puisqu'elles ont pour objectif de permettre de « décrire des objets individuels ou instances en les rangeant dans les classes et en spécifiant les valeurs des propriétés pertinentes dans leur cas. »¹¹⁵ C'est notamment le vocabulaire OWL, également présent dans le schéma vu précédemment, qui va offrir des moyens pour « déclarer équivalentes des classes ou des propriétés, et de tenir pour identiques des individus ; ce qui est précieux pour assurer l'interopérabilité des applications. »¹¹⁶ Il s'agit, en définitive, à travers l'utilisation d'ontologies, d'établir la structuration conceptuelle d'un domaine d'application avant de permettre de lier les représentations qui lui sont propres à des systèmes avec lesquels des alignements peuvent représenter un intérêt.

Puisque la réutilisation et l'interopérabilité sont les mots-clés du web des données, il n'est pas nécessaire de concevoir, ex nihilo, une ontologie à chaque fois que l'on veut décrire un type de données, mais il s'agit, du moins dans un premier temps, de « faire son marché »¹¹⁷, comme le dit Thomas Francart, fondateur de la société Sparna, parmi l'ensemble des vocabulaires déjà conçus et intégrés dans cette architecture globale. Ainsi, les principaux modèles qui s'inscrivent dans une réutilisation à grande échelle sont DCterms, FOAF, et SKOS, proposant respectivement des éléments de description générale, d'identification des personnes et des outils de conception de vocabulaires contrôlés. Toujours selon Thomas Francart, une telle réutilisation permet alors de

113 Ibid. Ceci étant, si dans sa feuille de route pour le web sémantique, Tim Berners-Lee rappelle que le modèle de base du RDF présente théoriquement de nombreuses possibilités, il signale qu'il ne constitue pas un outil en soi mais plutôt un ensemble d'assertions et de références permettant de « mapper », c'est-à-dire de mettre en correspondance, les données dans de nombreux formats et schémas.

114 Ibid.

115 Ibid.

116 Ibid.

117 FRANCART, Thomas. Penser, modéliser (pour le web de données) – Part 2/2. *Anaphore* [en ligne]. 4 juin 2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/penser-modeliser-pour-le-web-de-donnees-part-2/>

« bénéficier de la réflexion et des bonnes pratiques de modélisation qui ont sédimenté dans ce modèle et de rendre les données ainsi exprimées plus facilement compréhensibles et compatibles avec d'autres applications. »¹¹⁸ L'ontologie SKOS est justement l'un des exemples les plus parlants d'une transition vers un web sémantique puisque les systèmes d'organisation des connaissances, le plus souvent des thésaurus et autres vocabulaires contrôlés, « contiennent souvent des données sur des centaines, voire des milliers, de concepts et peuvent permettre aux institutions qui possèdent ces référentiels d'expérimenter à moindre coût la technologie prometteuse du web sémantique »¹¹⁹. L'ontologie SKOS, « qui se veut à la fois simple et compatible avec une majorité d'approches existantes (thésaurus, classifications, etc.) »¹²⁰ encourage la mise en relation de référentiels de natures différentes à l'aide de diverses propriétés dédiées à l'alignement des concepts. Ainsi, c'est par ce biais-là que de nombreuses institutions et organisations, notamment dans le secteur culturel, ont décidé de s'intéresser aux technologies du web sémantique.

Web des données et techniques documentaires, quelle imbrication ?

Nous le voyons, le recours à des vocabulaires structurés et à des schémas de métadonnées visant à décrire des ressources occupe une place centrale dans cette vision du web. La mise en œuvre de vocabulaires, notamment par l'intermédiaire de l'ontologie SKOS, mais également la mise en place de référentiels d'autorité au sein desquels les ressources sont identifiées à l'aide d'URI et décrites à l'aide de métadonnées en XML, constituent les principaux vecteurs de la mise en relation des données. En raison de cette imbrication, « le mouvement général du web sémantique mobilise des compétences parfaitement intégrées par les documentalistes. »¹²¹

En effet, les techniques documentaires, à l'ère de la redocumentarisation, où il s'agit de recomposer des ressources à partir de données fragmentées et éparées, notamment les compétences liées à la description et à l'indexation, qui sont parfaitement inscrites dans les pratiques des archivistes, seraient alors davantage qu'un outil mais plutôt un véritable pré-requis de la mise en œuvre du web des données :

118 Ibid.

119 DALBIN, Sylvie, BERMÈS, Emmanuelle, ISAAC, Antoine, WENZ, Romain, NICOLAS, Yann, MERABTI, Tayeb, ANGJELI, Anila, FRANCCART, Thomas, ROZAT, Lise, VANDENBUSSCHE, Pierre-Yves, VATANT, Bernard, RAIMOND, Yves et COTTE, Dominique. *Approches documentaires : priorité aux contenus. Documentaliste-Sciences de l'Information*, 31 janvier 2012, Vol. 48, n° 4, pp. 42-59.

120 Ibid.

121 Ibid.

« À notre sens, on ne peut aujourd’hui éditer de l’information sur le Web sans faire appel à des pratiques documentaires [...] Au fondement des métiers de la gestion de l’information se retrouvent les besoins suivants, quels que soient les outils et les environnements techniques : structurer, normaliser, qualifier l’information. Le web sémantique, [...] non seulement n’échappe pas à cette règle, mais il vient au contraire remettre sur le devant de la scène ce type de compétences que les évolutions du Web [...] avaient un temps semblé reléguer à l’arrière-plan. »¹²²

La construction de thésaurus et de langages documentaires divers, ainsi que la mise en place de référentiels sont alors au cœur de la construction des interfaces qui matérialisent la vision d’un web des données, et coïncident par la même occasion avec les pratiques des professionnels de l’information de tous bords. Dans le secteur culturel, l’exemple de l’application GINCO (Gestion Informatisée de Nomenclatures Collaboratives et Ouvertes), développée par le ministère de la Culture, symbolise cette orientation. Couche importante du programme HADOC¹²³, que nous évoquerons dans la partie qui suit, il s’agit justement d’un outil permettant « la gestion de thésaurus et de vocabulaires contrôlés [et] capable d’exporter ceux-ci en format SKOS. »¹²⁴

III.2. Web des données et archivage : quel potentiel pour la protection et la valorisation d’un patrimoine

Le web des données, un enjeu perçu à l’échelle de l’ensemble du secteur culturel

Dans le domaine de la culture, ces nouveaux enjeux liés aux évolutions du web imposent une nouvelle manière de concevoir la production documentaire et la communication institutionnelle. En effet, le web n’est alors plus simplement considéré comme « un simple média » mais doit être « conçu comme un cadre mondial d’échange, d’interopérabilité et d’interconnexion des données. »¹²⁵ C’est dans cette perspective que le Ministère de la culture a établi, en 2014, une feuille de route¹²⁶ proposant neuf axes stratégiques liés au web des données et dans laquelle une

122 Ibid.

123 HARmonisation des DONnées Culturelles

124 L’application GINCO. *Ministère de la Culture et de la Communication* [en ligne]. [Consulté le 7 août 2017]. Disponible à l’adresse : <http://www.culturecommunication.gouv.fr/Divers/Harmonisation-des-donnees-culturelles/Referentiels/Les-vocabulaires-scientifiques-et-techniques/L-application-GINCO>

125 SAJUS, Bertrand et LEROI, Marie-Véronique. Le développement du web des données culturelles. *ID – Information, données & documents*, février 2016, Vol. 53, n° 2, pp. 46-47.

126 Feuille de route – Métadonnées culturelles et transition Web 3.0 [en ligne] [Consulté le 19 avril 2017]. Disponible à l’adresse : http://www.culturecommunication.gouv.fr/var/culture/storage/pub/feuille_de_route__metadonnees_culturelles_et_transition_web_3_0_janvier_2014/index.htm.

vingtaine d'actions-clés ont été définies.¹²⁷ Dans ce contexte, un certain nombre de projets s'articule depuis plusieurs années déjà autour des principes du web des données, notamment dans les bibliothèques et musées où les projets Data BnF¹²⁸ et JocondeLab¹²⁹ ont bénéficié d'un rayonnement considérable à l'échelle nationale.

En raison de la manière dont sont structurés les catalogues de bibliothèques ou les bases de données muséales, mais également, puisque c'est là notre objet d'étude, les inventaires conçus par les archivistes, les données qui y sont exposées peuvent bénéficier des apports du web des données. Il s'agirait en effet d'offrir aux usagers de ces différents services « la possibilité de naviguer d'un site à un autre sans avoir connaissance des moyens techniques utilisés pour publier les données, ni même avoir conscience des ruptures ou des frontières entre chacun des sites. »¹³⁰ L'objectif clairement énoncé, et prégnant dans l'ensemble du domaine culturel, est celui d'une convergence des données culturelles dans l'optique d'une exposition plus efficace dans l'écosystème du web. Il s'agit alors de s'extirper des carcans qu'imposent les sites web et portails documentaires et archivistiques, qui constituent encore souvent des silos d'information. Pour Claire Sibille de Grimoüard le principal enjeu est donc de « mieux répondre aux attentes des utilisateurs en leur offrant un accès homogène aux différentes ressources patrimoniales, qu'elles soient archivistiques, bibliographiques ou muséales. »¹³¹ C'est justement cette intuition selon laquelle « les données d'autorité, les catalogues de bibliothèques disposent de référentiels pivots correspondant parfaitement à la logique et à la structure du web de données »¹³² qui a amené la BnF à développer son projet Data BnF.

127 Parmi celles-ci peuvent notamment être citées la mise en œuvre d'identifiants uniques et pérennes pour les ressources culturelles, l'interconnexion inter-institutionnelle des données culturelles ou encore le développement d'un dispositif de traçabilité des données numériques.

128 <http://data.bnf.fr/>

129 <http://jocondelab.iri-research.org/jocondelab>

130 DALBIN et al. art. cit. p. 39

131 GRIMOÛARD, Claire Sibille-de. Élaborer des normes de description... et les confronter à la pratique d'aujourd'hui. *La Gazette des archives*. 2012. Vol. 226, n° 2, pp. 165-177. DOI 10.3406/gazar.2012.4906.

132 VILLEMINOZ, Jérôme, [2016]. Le web de données à la BnF. In : *ABES : Arabesques* [en ligne]. n° 83. [Consulté le 2 mai 2017]. Disponible à l'adresse : <http://www.abes.fr/Publications-Evenements/Arabesques/Arabesques-n-83>. Comme le rappelle Jérôme Vileminoz, le choix qui a été fait dans le cadre de ce portail national celui d'une double réutilisation, avec des pages web consultables par les internautes et des jeux de données en RDF exploitables par les machines. C'est notamment la mise à disposition des vedettes matière Rameau en SKOS qui a favorisé la réutilisation des données et l'appropriation de la plateforme par les internautes, dans l'optique « de faire jouer à plein les ressorts du web de données. »

C'est également dans cette logique de réutilisation accrue des données que s'inscrit le programme HADOC du ministère de la Culture, au sein duquel l'harmonisation des données constitue le principal chantier, dans un souci « d'interopérabilité sémantique et organisationnelle et afin d'offrir une vision unifiée et cohérente des ressources culturelles. »¹³³ Dans le monde des Archives publiques, c'est sous la forme du portail national FranceArchives¹³⁴, lancé en mars 2017, que ces impulsions se sont matérialisées. Fruit du travail du SIAF et de nombreux services d'archives, il a pour objectif la valorisation des actions culturelles et scientifiques de l'ensemble des services d'archives français et le référencement centralisé sur le web des nombreux instruments de recherche qui découlent de leur activité. Pensé comme un véritable portail d'information et de consultation à destination des publics comme des archivistes, il s'agit d'un dispositif qui s'inscrit dans le mouvement de l'ouverture des données publiques. Ainsi, il s'agira pour nous de déterminer dans quelle mesure les services d'archives s'inscrivent, ou peuvent s'inscrire, dans de tels dispositifs et bénéficier eux aussi des technologies du web sémantique.

L'application au domaine des archives : spécificités et perspectives

Le développement des sites web d'archives et d'interfaces de consultation dédiées à l'accès aux documents numérisés symbolisent l'importance qu'accordent désormais les archives à l'ouverture et à la communication des fonds. En effet, depuis de nombreuses années déjà, Internet constitue le « le point de départ de la plupart des recherches »¹³⁵. Comme le rappelle Lorraine Filippozzi, archiviste suisse, la diffusion est au cœur des fonctions archivistiques, car elle « constitue l'aboutissement des activités de création et/ou d'acquisition, d'évaluation, de classification, de description et de préservation. »¹³⁶ Le web, et c'est d'autant plus vrai avec l'avènement du web des données qui a pour maîtres mots ouverture et interconnexion des données, présente

133 BRIATTE, Katell, 2014. Hadoc, un programme pour harmoniser les données culturelles. In : *Documentaliste-Sciences de l'Information*. 4 juillet 2014. Vol. 51, n° 2, p. 56-67. L'objectif du dispositif, dans son ensemble, est la consolidation des données d'identification des biens culturels au sein d'une « carte d'identité » unique ainsi que la mise au point d'un référentiel visant à apporter « les garanties d'authenticité, d'unicité, de qualité, d'exhaustivité et de pérennité attendues des producteurs comme des usagers. »

134 <https://francearchives.fr/>

135 FILIPPOZZI, Lorraine. Le web comme outil de diffusion des archives. *Ressi* [en ligne]. avril 2008. [Consulté le 17 avril 2017]. Disponible à l'adresse : http://www.ressi.ch/num07/article_044

136 Ibid.

alors un potentiel considérable pour les services d'archives. Dans le monde des archives, le recours aux technologies du web sémantique passe également par le biais d'une exposition de vocabulaires contrôlés. Dès 2009, les Archives de France ont mis en œuvre une réflexion concernant la révision du thésaurus pour l'indexation des archives locales, afin « d'expérimenter les opportunités offertes par le nouveau standard de structuration SKOS et le nouveau mode de publication proposés par le consortium W3C. »¹³⁷ L'alignement des référentiels, véritable moyen de créer des passerelles entre les différents domaines, est également un enjeu perçu par le SIAF, dans la mesure où une réflexion a été menée afin d'assurer la permanence des identifiants propres aux concepts de ce thésaurus, dans la perspective d'une mise en relation avec les autorités Rameau de la BnF ainsi que DBpedia.

Pour autant, si des initiatives portant sur l'alignement des thésaurus et sur la manière dont sont indexés les archives ont déjà porté leurs fruits, la manière dont les documents sont décrits, comme nous l'évoquions, pose la question de la compatibilité de la pratique avec les objectifs du web des données. D'après Karen Gracy, les limites inhérentes aux descriptions issues d'un encodage en EAD constituent un facteur limitant pour l'interopérabilité avec le modèle RDF. En effet, avec l'EAD, c'est l'aspect "narratif" et structuré de l'instrument de recherche qui est privilégié, reléguant ainsi les éléments individuels du document, de la pièce, au second plan et ne proposant alors pas nécessairement la granularité requise par le web sémantique.¹³⁸ Il s'agirait alors d'une véritable barrière à l'exposition des données archivistiques dans le web des données. De la même manière, les normes de description archivistique, conçues dans le respect du principe de respect des fonds et selon la description à plusieurs niveaux, correspondent à un modèle de description mono-entité. C'est notamment le cas de la norme ISAD(G) et par extension de l'EAD, qui dépendent strictement de la structure des instruments de recherche, contrairement à la norme ISAAR(CPF).¹³⁹

En outre, les problématiques liées à l'encodage des éléments de description des

137 SIBILLE, Claire. Archives et Données liées : nos outils sont-ils prêts à compléter la Toile ? *Modernisation et archives* [en ligne]. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://siaf.hypotheses.org/344>

138 GRACY, Karen F. Archival description and linked data : a preliminary study of opportunities and implementation challenges. *Archival Science*, 1 septembre 2015, Vol. 15, n° 3, pp. 239-294. DOI 10.1007/s10502-014-9216-2.

archives ne constituent pas la seule et unique difficulté, puisque des doutes ont longtemps subsisté sur la capacité des technologies du web sémantique à restituer l'ensemble des éléments contextuels et historiques indispensables en archivistique. Comme le rappelle Gautier Poupeau, data architect à l'INA, « l'absence de contextualisation est le principal reproche adressé à RDF par ses adversaires et il est vrai qu'au premier abord cela constitue son principal défaut. »¹⁴⁰ Or, cette contextualisation est absolument indispensable pour les services d'archives, étant donné que « l'archiviste décrit toujours une entité (comme une société par exemple) en regardant dans le rétroviseur »¹⁴¹ Il s'agit donc d'une réelle problématique pour la mise en conformité des données archivistiques avec les principes du web sémantique, puisque les notions de contexte et d'évolution sont évidemment parmi les éléments fondamentaux de la description des documents d'archives. En effet, un seul et unique URI ne permet alors pas de décrire l'évolution d'une entité. Pour répondre à cette problématique, le recours à des ontologies spécifiques est tout indiqué. Ainsi, comme le rappelle Thomas Francart¹⁴², un modèle tel que PROV-O, proposé par un groupe de travail du W3C, permet de satisfaire ces besoins de description de l'origine et de l'historique d'une entité. En intégrant, au sein même de l'ontologie, les notions d'entité et de spécialisation d'entité, l'ontologie PROV-O permet d'assurer le suivi des évolutions d'une ressource.

Malgré les difficultés qui caractérisent, la nécessité de « donner d'abord à nos outils logiciels une orientation pertinente et à même d'anticiper des évolutions technologiques, qui, pour l'instant, semblent tarder à recevoir toute l'attention qu'elles méritent »¹⁴³ et la participation « à l'adoption des technologies du web des données

139 En effet, comme le rappelle Louis Colombani, « le succès de la norme ISAD(G) et de la DTD EAD et, à l'inverse, la faible mise en pratique d'ISAAR et d'EAC témoignent de la prédominance du modèle mono-entité. »¹ Pour expliquer cette orientation qui semble bien ancrée au sein des services d'archives, il précise : « c'est probablement une question de culture ; mais, c'est aussi l'inadaptation des outils disponibles – en particulier le langage XML – à la mise en œuvre d'un modèle multi-entités qui a constitué un frein à l'évolution du modèle. »

140 POUPEAU, Gautier. Quel événement !? Ou comment contextualiser le triplet. *Les petites cases* [en ligne]. 2010. [Consulté le 21 avril 2017]. Disponible à l'adresse : <http://www.lespetitescases.net/quel-evenement-ou-comment-contextualiser-le-triplet>

141 FRANCAERT, Thomas. Penser, modéliser (pour le web de données) – Part 2/2. *Anaphore* [en ligne]. 4 juin 2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/penser-modeliser-pour-le-web-de-donnees-part-2/>

142 Ibid.

dans les archives »¹⁴⁴ sont des objectifs sur le long terme qui font sens pour les professionnels des archives. Ainsi, pour Myriam Pauillac, co-dirigeante de la société Anaphore, « sans remettre en cause les fondements de notre métier, la prise en compte des technologies du web sémantique inscrira sans aucun doute la pratique archivistique dans le changement du paradigme numérique de ce début de 21e siècle. »¹⁴⁵ L'absence d'un modèle conceptuel ou d'une ontologie de référence dans le monde des archives a pu alors paraître préjudiciable à cette ouverture vers le web des données. Par rapport aux bibliothèques et aux musées, les archivistes sont des « arrivants tardifs » qui « ont prudemment cherché leur voie. »¹⁴⁶ Si une adaptation aux modèles existants dans les autres domaines patrimoniaux, notamment au FrBr et au CIDOC-CRM, a de prime abord été envisagée, la communauté internationale des archivistes a finalement opté pour la mise au point d'un modèle conceptuel à part entière.

« Records in Context », un modèle conceptuel et une ontologie dédiées à la description archivistique

Le besoin d'un nouveau modèle de description archivistique repose sur le constat selon lequel les différentes normes préexistantes comportent un certain nombre d'incohérences, que celles-ci n'ont pas fait l'objet d'une articulation pensée dès le départ et qu'elles ne sont pas toutes ancrées dans les usages. L'objectif d'un tel modèle est au final de pouvoir « structurer de telles descriptions, [...] pour améliorer l'accès et la gestion de l'information, et en même temps [...] de rester compatible avec les données telles qu'elles sont aujourd'hui. »¹⁴⁷ Cependant, comme le signale Louis Colombani, « identifier les éléments de description ne suffit pas. Il s'agit de créer un véritable réseau de relations entre les différents types d'entités archivistiques. »¹⁴⁸ C'est dans cette optique qu'a été lancé le projet Records in Context, dans lequel le groupe de

143 PAUILLAC, Myriam. Web des données et archives – quel intérêt ? *Anaphore* [en ligne]. 23 novembre 2014. [Consulté le 15 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/web-des-donnees-et-archives-quel-interet/>

144 Ibid.

145 Ibid.

146 ANGJELI, Anila. Bibliothèques, archives, identifiants : les enjeux de la normalisation. *Hermès, La Revue*, 25 novembre 2013, n° 66, pp. 167-172.

147 FRAN CART. art. cit. p. 44

148 COLOMBANI, Louis. Modèle conceptuel et description archivistique ? *Anaphore* [en ligne]. 10 février 2015. [Consulté le 17 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/modele-conceptuel-et-archives/>

travail EGAD¹⁴⁹ de l'ICA est impliqué depuis 2012.¹⁵⁰ Le modèle conceptuel du même nom constitue la première facette de ce projet et se présente sous la forme d'un « document textuel enrichi avec des représentations graphiques. »¹⁵¹ Celui-ci, dans une optique de rationalisation des pratiques, « servira de référence dans les discussions entre professionnels » et devra permettre « aux archivistes de mieux dialoguer avec d'autres professionnels de l'information sur les concepts et les approches métier. »¹⁵² Il s'agit alors, d'après une logique influencée par les évolutions technologiques et les orientations du web des données, notamment des vocabulaires qui permettent son fonctionnement, de faire évoluer la description à plusieurs niveaux, héritée du corpus normatif de l'archivistique, vers un modèle qui échapperait alors aux limites imposées par la structure des instruments de recherche. Pour ce faire, RiC-CM établit la distinction entre un document archivistique (Record), un ensemble de documents (RecordSet), et un élément d'un document d'archives (Record Component). Il permet de définir les différentes entités de l'archivage, d'en préciser les propriétés, tout en indiquant l'ensemble des relations entre ces entités.

L'ontologie RiC (ou RiC-O), est « l'expression informatisée (machine-readable) du Modèle conceptuel, en langage W3C OWL [et] fournira le cadre sémantique pour exprimer les concepts archivistiques, les différents composants des collections. »¹⁵³ Par ailleurs, dans l'optique d'une compatibilité avec l'existant et pour demeurer ouvert aux évolutions futures, le modèle n'a pas vocation à se subsister aux différents schémas de métadonnées et langages évoqués jusque-là, mais plutôt de fournir un cadre de haut niveau permettant d'harmoniser la pratique de la description archivistique. Le principal intérêt d'une telle ontologie résiderait alors dans sa capacité à outrepasser la conception « monolithique », en application actuellement dans le domaine de la description archivistique. Ainsi, le dispositif RiC,

149 Groupe d'experts sur la description archivistique

150 ANGJELI, Anila. Congrès 2015 de la SAA. Episode 2 : les Archives dans leur contexte (Records in context, RiC). Modernisation et archives [en ligne]. 5 novembre 2015. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://siaf.hypotheses.org/493> « Les travaux se fondent d'une part sur les principes archivistiques, notamment celui de Provenance et celui de Respect de l'ordre originel (ou Respect des fonds) et d'autre part, sur la séparation progressive des composantes de la description archivistique, évolution reflétée par l'élaboration successive des quatre normes, ISAD(G), ISAAR(CPF), ISDF et ISDIAH. »

151 Ibid.

152 Ibid.

153 Ibid.

tel qu'il a été pensé par EGAD, propose une vision ouverte de la description d'archives, ce qui laisse penser qu'il s'agit là également d'anticiper la mise en correspondance de concepts archivistiques avec ceux d'autres domaines culturels. Dans la mesure où l'un des objectifs d'une ontologie est de permettre l'interopérabilité avec d'autres modèles similaires, l'intérêt d'une éventuelle mise en relation avec les autres modèles conceptuels pensés pour décrire les ressources ou objets culturels (notamment le FrBr et CIDOC-CRM) est incontestable et constitue l'un des enjeux principaux d'un projet tel que celui-ci. Claire Sibille de Grimoüard met ainsi en évidence l'apport d'une telle ontologie en matière d'accès aux ressources¹⁵⁴, ce qui laisse supposer une interconnexion pensée avant tout pour être au service des publics, et par conséquent, un objectif d'ouverture des données culturelles. Ceci étant, l'efficacité du modèle conceptuel et de l'ontologie qui l'accompagne semble, avant même leur mise en application, faire encore l'objet d'interrogations¹⁵⁵, ce qui pose la question de l'impact qu'un tel dispositif peut avoir sur des pratiques profondément ancrées pour les professionnels concernés.

III.3. Le web des données, entre enjeux de gouvernance et stratégie de diffusion

Description archivistique et données liées, entre antagonisme et complémentarité

Nous l'avons vu, les spécificités liées au traitement des archives, à savoir la description à plusieurs niveaux, la hiérarchie et les règles en matière d'héritage qu'elle implique, qui découlent directement des principes de contextualisation et de respect des fonds, peuvent constituer des obstacles à la mise en application des technologies du web sémantique dans le domaine des archives. Ces freins en termes de compatibilité de ces deux manières de penser la description archivistique sont considérés comme un

154 GRIMOÛARD, Claire Sibille de. Vers un modèle conceptuel international pour la description archivistique. *Arbido* [en ligne], 2012, Vol. 1. [Consulté le 17 avril 2017]. Disponible à l'adresse : http://www.ica.org/sites/default/files/EGAD_French.pdf

155 Cf. InterPARES Trust responds to EGAD-RiC. In : *InterPARES Trust* [en ligne]. 11 décembre 2016. [Consulté le 17 avril 2017]. Disponible à l'adresse : <https://interparestrust.com/2016/12/11/inter pares-trust-responds-to-egad-ric/>. Le rapport publié par le groupe de travail InterPARES en décembre 2016 fait état de leurs inquiétudes vis-à-vis du processus de conception de RiC et de sa capacité à subvenir aux besoins en matière de description et d'interopérabilité des différents services d'archives. Les commentaires d'InterPARES soulignent la difficulté à aboutir à un modèle universel permettant de satisfaire les besoins de tous les services d'archives.

véritable défi pour l'intégration des données archivistiques dans le web des données. C'est notamment la question de la granularité de l'information et de l'adéquation des standards d'encodage existants, mais également des systèmes d'information archivistique, qui sont au cœur des réflexions menées par les archivistes et groupes de travail qui se sont penchés sur ces aspects. Pour expliquer la prévalence d'un modèle de description mono-entité pour décrire les documents d'archives, Louis Colombani soupçonnait déjà l'impact de facteurs « culturels. »¹⁵⁶ Dès lors, nous avons pour hypothèse que la culture professionnelle des archivistes, les formations qu'ils ont pu suivre et le rapport particulier qu'ils entretiennent avec l'objet document, à travers une culture de l'imprimé très développée mais aussi via la création d'instruments de recherche pensés pour le papier, peuvent en partie expliquer leurs difficultés à percevoir le potentiel du web des données et à mettre en œuvre un modèle qui satisfasse ce pré-requis de l'interopérabilité.

Le recours aux technologies du web sémantique au service de la gouvernance de l'information ?

Au-delà des enjeux technologiques et de la nécessité de faire coïncider les outils du web sémantique avec les pratiques des archivistes, il s'agit également de s'intéresser à la portée fonctionnelle des vocabulaires et référentiels et à ce qu'ils impliquent en termes d'interactions interprofessionnelles, au regard de l'objectif d'interopérabilité qui les caractérise. Sylvie Dalbin précise en effet que « ce travail de portage aux couleurs du web de données est systématiquement l'occasion d'une révision des vocabulaires utilisés dans les différents systèmes et de leur reconfiguration en tenant compte de cette nouvelle orientation de mutualisation et de partage. »¹⁵⁷ Il s'agit alors pour chaque catégorie de professionnels concernée d'y piocher les éléments nécessaires, en fonction de « ses caractéristiques métiers et de ses exigences fonctionnelles. »¹⁵⁸ De ce fait, les outils découlant de cette volonté d'intégrer les données de l'archivage au web des données peuvent rejoindre les problématiques propres à la gouvernance de l'information. En outre, comme nous l'évoquons, la

156 COLOMBANI, art. cit. p. 45

157 DALBIN, Sylvie, BERMÈS, Emmanuelle, ISAAC, Antoine, WENZ, Romain, NICOLAS, Yann, MERABTI, Tayeb, ANGJELI, Anila, FRANCCART, Thomas, ROZAT, Lise, VANDENBUSSCHE, Pierre-Yves, VATANT, Bernard, RAIMOND, Yves et COTTE, Dominique. *Approches documentaires : priorité aux contenus. Documentaliste-Sciences de l'Information*, 31 janvier 2012, Vol. 48, n° 4, pp. 42-59.

158 Ibid.

transversalité accrue de la GI impose aux archivistes de développer de nouvelles formes d'interactions avec d'autres typologies de professionnels afin d'être en mesure de dialoguer et de faire valoir leurs compétences. Dans cette hypothèse, les échanges entre les différentes catégories de professionnels, et notamment entre archivistes et informaticiens, qui ne partagent ni le même vocabulaire ni la même vision de l'archivage, semblent alors être cruciaux pour développer un système d'archivage électronique robuste et pérenne.

Dans la mesure où la GI et le web des données reposent sur des outils similaires, à savoir des vocabulaires et des référentiels communs ainsi que sur une volonté d'harmoniser les données, nous pouvons émettre l'hypothèse selon laquelle le recours à des identifiants uniques et pérennes ou encore la conception de vocabulaires en SKOS seraient envisagés, dans un premier temps, dans l'optique d'une rationalisation des pratiques. En outre, puisque l'archivage électronique constitue un enjeu majeur pour les organisations qui veulent mettre en place une gouvernance de l'information, et qu'il doit répondre à des besoins de fiabilité et de traçabilité, les technologies du web sémantique pourraient, dans une certaine mesure, permettre de répondre à certains de ces enjeux. En effet, le recours à des ontologies qui permettent d'assurer la contextualisation des données et de retracer leur historique pourraient permettre de garantir la valeur probante d'un document d'archives. Pour autant, nous avons également pour hypothèse que les archivistes peuvent se sentir déboussolés et pourraient avoir à réévaluer leur positionnement, par rapport à la chaîne de traitement du document et de la donnée, mais également vis-à-vis des autres professions qui gravitent autour de ces enjeux. Il s'agira alors pour nous d'interroger la manière dont les archivistes perçoivent ces mutations et de déterminer si leurs modes opératoires actuels peuvent coïncider avec l'ensemble des normes, outils et processus qui se développent dans ce contexte, et dans quelle mesure ces derniers peuvent bénéficier des apports du web sémantique.

La diffusion des données archivistiques sur le web, quelle réalité pour les usagers ?

À travers la normalisation des pratiques, l'adoption de standards au service de l'interopérabilité entre les institutions patrimoniales et la mise au point d'outils de

recherche adaptés, ce sont également les publics des archives qui vont se placer en première ligne de ces mutations. C'est alors l'efficacité de dispositifs actuels qui est remise en question : « les approches traditionnelles de l'interopérabilité ont montré leurs limites : la constitution de portails fournissant un accès fédéré à plusieurs bases de données présente le double inconvénient d'offrir une expérience de recherche appauvrie aux usagers, et de les obliger à se connecter sur le portail pour faire une recherche, ce qui suppose donc de connaître *a priori* son existence. »¹⁵⁹ L'intégration des données culturelles au sein du web des données constituerait alors une réelle révolution en termes d'usage, puisqu'il offrirait aux internautes « la possibilité de naviguer d'un site à un autre sans avoir connaissance des moyens techniques utilisés pour publier les données, ni même avoir conscience des ruptures ou des frontières entre chacun des sites. »¹⁶⁰

La prise en compte des usages est d'autant plus importante que certaines réserves ont pu être émises dernièrement au sujet de la manière dont les professionnels du secteur culturel s'approprient la question du web des données. En effet, Jonathan Rochkind, ingénieur à l'Université John Hopkins de Baltimore, dans son billet de blog *Linked Data Caution*¹⁶¹, met en garde les bibliothécaires, mais plus largement tous les professionnels de l'information qui sont concernés par ces évolutions technologiques. C'est l'engouement massif pour ces technologies à partir de la fin des années 2000 qui le pousse à s'interroger de la sorte. Dans ce billet critique où prudence est le maître mot, il invite les professionnels à un certain scepticisme et à placer la focale sur les usagers (« stay user-centered »), en rappelant que l'utilisation des technologies du web sémantique ne doit pas être une finalité en soi mais plutôt un moyen d'améliorer les services proposés. Les principales difficultés à mettre en œuvre des technologies seraient selon lui avant tout liées à l'interopérabilité des données et à l'utilisation de modèles et référentiels communs, mais aussi à la réticence de certaines structures à partager leurs données. Dès lors, puisqu'il est encore difficile, en dehors de quelques projets qui ont su trouvé un certain rayonnement dans le secteur culturel, de prendre

159 DALBIN et al. art. cit. p. 48

160 Ibid.

161 ROCHKIND, Jonathan, 2015. *Linked Data Caution*. In : *Bibliographic Wilderness* [en ligne]. 23 novembre 2015. [Consulté le 31 juillet 2017]. Disponible à l'adresse : <https://bibwild.wordpress.com/2015/11/23/linked-data-caution/>.

la réelle mesure du potentiel du web des données, et que la mise en application des technologies du web sémantique dans le domaine des archives n'arrive que tardivement, les enjeux d'un éclatement des silos d'information et de l'ouverture des données à travers leur exposition massive au sein de gisements dédiés pourraient encore être difficiles à envisager pour les services d'archives. Par extension, le potentiel de ces technologies et leur intérêt pour les publics des archives, dont la typologie (chercheurs, historiens, généalogistes, etc.) peut également jouer un rôle, semble difficile à évaluer. Nous l'avons vu, les outils (instruments de recherche, portails, moteurs de recherche dédiés, etc.) à destination des publics vont alors indéniablement être amenés à évoluer mais un manque de visibilité semble encore régner sur les orientations de ces mutations. Cette interrogation est d'autant plus importante que les dispositifs qui émergent de cette conception du web impliquent « des logiques d'accès et de navigation spécifiques. »¹⁶² Il s'agit alors de trouver les moyens de valoriser ces nouvelles plateformes et de faire le lien entre les usages actuels et ceux qui vont découler du web des données. Dès lors, comme le rappelle Romain Wenz, « ces jeux de données méritent une médiation [et] cette médiation passera d'ailleurs par l'évolution de l'interface »¹⁶³, dans l'optique de la création « d'un véritable service aux utilisateurs, au-delà de la seule mise à disposition d'une infrastructure de récupération et d'interrogation. »¹⁶⁴ Nous veillerons donc à accorder une attention particulière à la manière dont les technologies du web sémantique peuvent contribuer à la diffusion des archives, tout en questionnant l'impact qu'elles peuvent avoir sur les comportements des usagers, en termes de navigation et de réutilisation.

Entre outils de gouvernance et potentiels vecteurs d'une meilleure diffusion du patrimoine culturel conservé par les services d'archives, les technologies du web sémantique représentent à la fois un défi et une opportunité pour les archivistes. C'est alors cette dualité qui s'est trouvée au cœur de notre questionnement, et qui a justifié la mise au point d'un dispositif d'expérimentation adapté, que nous présenterons dans la partie qui suit.

162 DALBIN et al. art. cit. p. 48

163 Ibid.

164 Ibid.

Deuxième partie : l'élaboration du protocole d'expérimentation

Afin d'être en mesure de répondre de manière optimale aux problématiques mises en lumière dans notre première partie et de pouvoir nous positionner par rapport aux hypothèses émises à cette occasion, il a fallu nous interroger sur la nature de l'expérimentation à mettre en œuvre. Dans la mesure où les éléments avancés dans la première partie se veulent le reflet théorique des évolutions auxquelles font face les services d'archives, cette étape de confrontation des hypothèses avec la réalité du terrain se devait d'inscrire ces dernières dans un cadre plus concret afin d'en tirer des conclusions représentatives des bouleversements auxquels les archivistes sont confrontés avec l'avènement de l'archivage électronique. La partie qui suit sera donc dédiée à l'explicitation des différentes étapes qui ont mené, d'une part, au choix du protocole d'expérimentation et d'autre part, à sa mise en œuvre et à la méthodologie appliquée. La présentation des modalités d'expérimentation se fera donc en deux temps, puisque nous avons opté pour un mode opératoire mixte, à savoir d'un côté une étude de cas des procédures du projet de Système d'Archivage Électronique Mutualisé girondin, en mettant l'accent sur la place des technologies du web sémantique au sein du système étudié, et de l'autre une approche qualitative visant à cibler la perception que les différents acteurs du projet, et plus particulièrement les archivistes, ont de ces évolutions. Nous nous attacherons également à présenter le contexte dans lequel la méthodologie d'expérimentation a pu être appliquée, afin de nous assurer que la troisième partie puisse être entièrement dédiée à l'analyse des résultats.

I. L'étude de cas à partir de l'exemple du SAEM girondin : attentes et objectifs

En prélude à la présentation de la méthodologie en question, il convient de signaler que l'ensemble de ce protocole a pu être mis en place dans le cadre d'un stage réalisé auprès du Conseil départemental de la Gironde, au cours duquel nous étions directement intégrés au comité projet, c'est-à-dire l'équipe opérationnelle composée

d'archivistes et d'informaticiens, du système d'archivage électronique mutualisé girondin. Dans un premier temps, nous expliquerons en quoi ce contexte d'expérimentation a pu constituer une opportunité pour observer les éléments soulevés en première partie et pourquoi nous avons privilégié cette démarche à d'autres approches que nous envisagions initialement.

I.1. Le choix de l'étude des procédures du SAEM : le contexte de l'expérimentation

Contexte et considérations pratiques de l'expérimentation

Le contexte que nous venons d'exposer a été déterminant pour établir les orientations à suivre pour l'expérimentation. En effet, si la documentation disponible en ligne offre effectivement une vision d'ensemble des différentes étapes du projet et des moyens mise en œuvre pour assurer le fonctionnement du SAEM, le fait d'être intégré au comité projet nous a permis d'observer au plus près les rouages du système et d'étudier les subtilités des interactions entre les différents acteurs ainsi que d'entrevoir les difficultés, qu'elles soient techniques ou fonctionnelles, qui peuvent émerger dans le cadre d'un tel dispositif. En effet, sur une période de trois mois et demi (du 11 avril au 28 juillet 2017), nous avons pu assister à des réunions hebdomadaires du groupe projet, aux échanges avec les prestataires ainsi qu'aux différentes présentations effectuées auprès des instances décisionnelles et à l'attention d'autres services d'archives départementaux. Conscients qu'une telle proximité peut introduire un biais lors de l'expérimentation, nous avons fait le choix de ne restituer que des éléments factuels à partir de cette étude de cas, laissant la parole aux professionnels impliqués et à la vision qui leur est propre à l'occasion d'entretiens dont nous détaillerons également par la suite les modalités. Pour autant, c'est justement cette proximité qui a permis de cibler les composantes du projet les plus à même de répondre à nos interrogations, et qui a constitué une plus-value par rapport à une stratégie de benchmarking que nous avons pu envisager au départ.

Un questionnaire initial : favoriser l'approche comparative ou l'étude de cas ?

Lorsque nous nous sommes posé la question de la méthodologie à adopter, plusieurs solutions, qui semblaient alors toutes parfaitement viables, s'offraient à nous. Ainsi,

nous avons longuement songé à favoriser une approche comparative reposant sur une enquête auprès de différents services d'archives, qui aurait eu pour objectif de mettre en lumière la place, ou l'absence, des technologies du web sémantique dans les logiciels, dans les interfaces de consultation et dans les pratiques des professionnels des archives. Ceci étant, pour garantir l'efficacité d'un tel dispositif, il aurait fallu être en mesure de repérer des interlocuteurs à la fois à l'aise avec les sujets que nous souhaitons aborder et représentatifs des impulsions qui peuvent avoir lieu dans les services publics d'archives. Dès lors, à partir de nos recherches et en raison de la place qu'occupent ces technologies dans la mise en place du projet SAEM, nous avons préféré nous orienter vers une analyse détaillée des procédures du projet, qui ont fait l'objet d'une documentation relativement exhaustive, et ce depuis la genèse du dispositif. En outre, et c'est ce que nous soulignerons dans la partie qui suit, le rayonnement du projet au niveau national ainsi que l'organisation des différentes équipes qui y travaillent, notamment en ce qui concerne la répartition des rôles et la méthodologie de gestion de projet retenue, nous ont semblé constituer une opportunité intéressante pour étudier l'ensemble des questions relatives à la gouvernance de l'information évoquées dans la partie théorique de ce mémoire.

I.2. Le système d'archivage électronique mutualisé girondin, un exemple représentatif ?

Un projet considéré comme pionnier dans le domaine de l'archivage électronique pour les collectivités

Initié dès l'année 2011, le projet de Système d'Archivage Électronique Mutualisé girondin a fait l'objet d'un long processus de réflexion et de conception mené par le Conseil départemental de la Gironde ainsi que les Archives départementales, qui sont sous sa tutelle, mais aussi par la ville de Bordeaux et par la communauté urbaine de Bordeaux (désormais Bordeaux Métropole). L'ancienne région Aquitaine était également impliquée, au lancement du projet. Les principaux objectifs du système d'archivage, tels qu'ils sont définis dans la politique d'archivage du projet, sont la réalisation des versements, la conservation pérenne et intègre des données, la gestion du cycle de vie des archives dans son ensemble ainsi que la communication des archives dans le respect des différentes procédures contractuelles. La volonté de

répondre aux impératifs liés à l'essor de la dématérialisation des documents produits par les administrations et aux enjeux liés aux besoins d'authenticité, d'intégrité et de pérennité de ces documents, constitue le point de départ du SAEM girondin. L'ensemble du projet s'inscrit donc dans la politique de gestion du cycle de vie des documents des collectivités impliquées, selon laquelle la gestion des flux documentaires doit être normalisée afin de garantir l'adaptation au cadre réglementaire et normatif dans ses évolutions futures, la capacité à produire des preuves de l'activité de l'administration mais également à garantir la fiabilité, la pérennité et l'accès de leur patrimoine d'information numérique. Le système a également pour but de permettre le stockage et l'administration des données, en archivage intermédiaire ou définitif et doit permettre la communication des métadonnées descriptives dans l'optique d'une restitution aux utilisateurs.

Afin de distinguer les principales étapes du projet et de donner aux instances décisionnelles une vision d'ensemble des développements prévus, le projet a été découpé en trois phases distinctes. Une première phase d'étude et de prototypage, qui s'est étalée d'avril 2012 à mai 2014, une deuxième phase de développements complémentaires, toujours en cours et prévue pour s'achever à la fin de l'année 2017, et une phase finale consistant en l'exploitation puis à la généralisation du SAEM, entre 2018 et 2019. De plus, à partir de 2019, l'ensemble du système doit faire l'objet d'une industrialisation et s'ouvrir à d'autres utilisateurs. Cette logique d'ouverture constitue par ailleurs l'un des éléments fondamentaux du projet, puisque le recours à des logiciels libres sur l'ensemble de la chaîne, permettant la mise à disposition des développements et la libre exploitation des systèmes prévus, a occupé une place centrale dès les phases d'étude et de prototypage. Ainsi, c'est la licence de réutilisation Creative Commons CC-BY-SA¹⁶⁵ qui a été retenue pour les études et la licence MIT¹⁶⁶ pour les développements. Afin d'être en mesure d'appréhender notre objet d'étude dans toute sa complexité et de comprendre les choix qui ont été faits quant au périmètre de l'expérimentation, il nous paraît également indispensable de présenter en amont les différentes briques applicatives, ou modules, qui sont au cœur du

165 Licence libre copyleft Creative Commons.

166 Licence libre non copyleft du Massachusetts Institute of Technology.

fonctionnement du SAEM girondin.¹⁶⁷ Ainsi, le système repose sur une interconnexion entre trois modules distincts, qui ont pour chacun d'entre eux des fonctionnalités et des objectifs clairement définis. Le module de gestion des processus d'archivage, que nous pourrions appeler « GED SAS » lorsque nous aurons à nous y référer, permet de gérer les versements, la consultation, la restitution ainsi que l'élimination de documents dans un espace documentaire dédié reposant sur une Gestion Électronique des Documents à partir du logiciel libre Alfresco.¹⁶⁸ Il est utilisé aussi bien par les services versants que par les archivistes. Le module de gestion de l'archivage et de la conservation, reposant sur la solution Asa@lae proposée par l'ADULLACT¹⁶⁹, a pour objectif la conservation pérenne des archives numériques. Il s'agit donc de la brique SAE du dispositif et il est utilisé essentiellement par les archivistes. En ce qui concerne le module de gestion des données de référence, que nous nommerons par la suite « référentiel », celui-ci permet de gérer les données relatives aux acteurs de l'archivage, aux vocabulaires contrôlés et aux profils de versement. Il s'agit d'un outil qui est avant tout à destination des archivistes, qui en sont les gestionnaires. Il convient également de souligner que ces modules peuvent fonctionner de manière indépendante et sont prévus de manière à pouvoir être implémentés dans d'autres systèmes. Pour autant, c'est la combinaison de ces différents éléments qui garantit la compatibilité avec l'ensemble des exigences fonctionnelles de l'archivage électronique et, ce faisant, qui permet de faciliter les tâches de gestion récurrentes qui incombent aux services versants et aux archivistes.

Si nous avons estimé que ce projet était l'occasion idéale pour observer les éléments mis en lumière dans la première partie de ce mémoire, ce n'est pas seulement en raison du contexte favorable à une expérimentation en profondeur, mais bel et bien parce que celui-ci présente une singularité essentielle par rapport aux autres SAE, en intégrant les technologies du web sémantique à travers le recours à un référentiel de données.

167 Cf. schéma présentant la vision modulaire du projet en annexe 9.

168 Système de gestion de contenu créé par Alfresco Software, sous licence libre.

169 Association des développeurs et utilisateurs de logiciels libres pour les administrations et les collectivités territoriales.

Le module référentiel : pierre angulaire de l'harmonisation des données et symbole du recours aux technologies du web sémantique

Le module de gestion des données de référence du SAEM girondin constitue en effet une particularité qui lui est propre et résulte d'une réflexion menée pour garantir l'interopérabilité avec les outils utilisés quotidiennement par les archivistes et dans la perspective d'une réutilisation des données. De ce fait, l'intégration du module référentiel au sein de l'architecture fonctionnelle du SAEM girondin répond à un objectif de centralisation de l'ensemble des données de référence dans une seule et unique brique applicative.¹⁷⁰ Afin de permettre cette son interopérabilité, c'est le protocole OAI-PMH¹⁷¹ qui a été choisi pour permettre les échanges de données, qui sont automatisés à l'aide d'un certain nombre de web services développés sur mesure. La GED SAS et le SAE sont alors en mesure de récupérer, depuis le référentiel, les données relatives aux profils SEDA, aux acteurs de l'archivage ainsi que les différents vocabulaires contrôlés et les concepts qui y sont rattachés afin de permettre l'indexation des versements et des objets d'archives qui les constituent. En ce qui concerne l'intégration de ces vocabulaires au sein du module, c'est l'ontologie SKOS, que nous avons évoqué en première partie, qui est mise en application et qui permet la conformité, là encore dans une optique d'interconnexion et de réutilisation, avec les standards du web des données.

Afin de mener à bien le développement de ce module et de garantir l'interfaçage avec les solutions retenues avant le lancement de la deuxième phase du projet, l'équipe du projet a fait appel au prestataire Logilab. Cette société a également mené les projets du portail Data BnF et du portail FranceArchives, qui jouissent tous deux d'un certain rayonnement au niveau national. C'est l'outil CubicWeb qui constitue le socle technique de ces différentes solutions, puisqu'il est articulé autour de « cubes »¹⁷² reposant sur des modèles de données et des ontologies OWL. En outre, l'outil permet, et c'est là l'une des attentes principales par rapport à la conception du référentiel, de relier les données entre elles grâce à des liens et des identifiants pérennes et uniques, en supprimant les ambiguïtés, mais également en décrivant l'ensemble de ces relations

170 Cf. schéma concernant la place du module référentiel en annexe 9.

171 Protocole international offrant un moyen d'échanger et de moissonner des métadonnées, notamment entre plusieurs institutions.

172 Composant logiciel reposant sur trois éléments : un modèle de données (sous la forme d'un schéma), une logique (des entités) et une interface utilisateur (des « vues »).

selon le modèle RDF. Le référentiel du SAEM girondin permet de créer, d'importer, de gérer, d'interconnecter et de consulter les données de référence produites par les archivistes, à savoir les acteurs de l'archivage¹⁷³, les notices d'autorité conformes à la norme ISAAR(CPF), les vocabulaires contrôlés, mais aussi les profils d'archivage SEDA, qui définissent les contraintes liées à un versement d'archives ainsi que les unités d'archives.¹⁷⁴ Son intérêt principal est alors de relier l'ensemble de ces entités employées pour effectuer des versements d'archives par des relations typées. A titre d'exemple, il est possible de décrire un producteur ou une unité d'archives à l'aide d'un concept issu d'un vocabulaire contrôlé ou bien de décrire les relations existantes entre des producteurs, qu'elles soient hiérarchiques, temporelles ou associatives.

Un mode de gouvernance particulier : l'organisation du comité projet et l'impact de la mutualisation

En raison de l'implication de plusieurs collectivités dans le projet et de l'intégration de professionnels d'horizons divers dans les différentes instances du projet, ce dernier reposant notamment sur des binômes informaticien-archiviste pour chacune des collectivités impliquées, nous avons également estimé qu'il s'agissait d'un terrain propice à l'observation de l'évolution des pratiques que peuvent engendrer ces modes de gestion de projet qui favorisent la transversalité. L'importance de la conception mutualisée du SAEM ainsi que la volonté de s'inscrire dans une politique plus large, comme le démontre le travail mené auprès des équipes du SIAF et les interactions avec les acteurs du programme interministériel VITAM¹⁷⁵, favorisent également l'analyse de la portée des impulsions liées à la conception de référentiels de données au niveau national.

En outre, le fait que le projet s'appuie, et ce depuis la rédaction du cahier des charges, sur les méthodes de gestion agiles¹⁷⁶, qui favorisent une certaine forme d'horizontalité, des interactions accrues entre les différents acteurs et avec les prestataires ainsi qu'un

173 C'est-à-dire les autorités administratives (collectivités), composées d'unités administratives (services versants, producteurs, ou service d'archives), ainsi que les agents, contacts référents de ces entités.

174 Une unité d'archive correspond à un dossier type, réutilisable au sein d'un profil. Il peut n'être composé que d'une pièce (un fichier) ou d'un ensemble de fichiers et peut lui-même être découpé en sous-dossiers.

175 Valeurs Immatérielles Transmises aux Archives pour Mémoire.

176 Méthodologie de gestion de projet issue du monde de l'informatique et apparue au tournant des années 2000. Cf. <http://agilemanifesto.org/iso/fr/manifesto.html>

fonctionnement itératif, nous a conforté dans nos choix. En effet, et nous aurons l'occasion d'y revenir, ces méthodes de gestion, si elles sont aujourd'hui très prégnantes parmi les informaticiens et dans les DSI, ne sont pas pour autant répandues dans les services d'archives. Cet aspect constitue donc une couche supplémentaire permettant d'observer les interactions interprofessionnelles et les enjeux en matière de conduite du changement.

I.3. L'approche suivie et ses différentes étapes

L'analyse de la production documentaire et des procédures du projet

Dans un premier temps, et afin d'avoir une vision d'ensemble des différentes briques de l'architecture technique du SAEM girondin, nous nous sommes familiarisés avec l'ensemble de la production documentaire du projet, accessible pour une partie sur le blog qui lui est dédié et en intégralité sur les différents espaces numériques de stockage que les différentes collectivités impliquées ont mutualisé afin de faciliter leurs échanges. Cette étape s'est avérée essentielle pour comprendre les mécanismes techniques du SAEM et percevoir les enjeux liés à l'imbrication des différentes couches applicatives, telle qu'elle a été envisagée dès la première phase du projet. Afin de réunir l'ensemble des éléments nécessaires à notre expérimentation, nous avons favorisé la diversité des sources d'information et avons choisi d'explorer l'ensemble des procédures accessibles. Ainsi, c'est la politique d'archivage (PA) ou politique du système d'archivage (PSA) du projet, qui a nous a permis d'analyser la vision initiale du projet et d'en comprendre les enjeux et objectifs. Il s'agit en effet d'un document stratégique, élaboré dès la première phase du projet, qui constitue le socle des différents choix effectués et un biais pertinent pour observer la gouvernance du projet dans son ensemble. Les spécifications fonctionnelles des différentes couches applicatives du projet, pour la plupart disponibles sur le blog du SAEM girondin¹⁷⁷, nous ont permis d'entrevoir les attentes fonctionnelles par rapport au système d'archivage dans son ensemble et de cibler les éléments en lien avec le potentiel du web sémantique. Les multiples conventions de partenariat ainsi que les notes d'intention en matière de répartition des rôles des différents acteurs du projet ont également constitué un gisement d'informations considérable pour analyser la gouvernance du projet. Le

177 <http://saem.e-bordeaux.org/>

cahier des charges ainsi que le cahier des clauses techniques particulières (CCTP), dans lesquels sont clairement énoncées les attentes initiales des collectivités à l’initiative du projet, ont eux aussi fait l’objet d’une attention particulière au cours de notre analyse. De plus, la dernière version du manuel utilisateur et les différents documents de communication et de présentation aux instances décisionnelles ou à d’autres services d’archives ont également alimenté notre réflexion. Enfin, ce sont les traces des échanges avec le SIAF et avec l’équipe du programme VITAM qui ont pu être utilisées pour parvenir à nos résultats, ainsi que certains comptes-rendus de réunions du comité projet et d’ateliers avec les différents prestataires.

Nous avons par ailleurs choisi de recentrer l’étude sur les points-clé du projet et les intentions ainsi que les objectifs de ceux qui y ont contribué. Dans la plupart des cas, une analyse poussée des dispositifs techniques qui permettent l’interconnexion des différentes briques, bien qu’il fût nécessaire de les connaître pour comprendre le fonctionnement du système, n’aurait pas généré suffisamment d’éléments pertinents relatifs à notre objet d’étude.

La manipulation des outils et l’étude des tickets et user stories

Notre intégration à l’équipe du projet ne se limitait évidemment pas à la participation aux réunions, puisque nous avons eu l’occasion de manipuler les différentes couches applicatives du SAEM, et plus particulièrement son module référentiel. De ce fait, cela nous a permis de visualiser, de manière très concrète, la façon dont sont intégrés les vocabulaires en SKOS, les notices d’autorité en EAC-CPF ainsi que la manière dont a été pensée l’interconnexion avec le SIA des Archives départementales de la Gironde.

Puisque nous étions chargés d’effectuer des tests utilisateurs sur les différents outils qui composent le système, et plus particulièrement sur le module référentiel, nous avons rapidement eu accès au système de gestion de tickets que Logilab a pu mettre en place pour assurer le suivi des développements et de la correction des anomalies. C’est sur cette plateforme que les membres du comité projet ont pu directement rédiger des « user stories »¹⁷⁸, qui, en raison de leur formulation normalisée, débutent systématiquement par « en tant qu’archiviste, je souhaite... » ou « en tant

178 Terminologie propre aux méthodes agiles, correspond à un récit utilisateur décrivant le contenu d’une fonctionnalité à développer.

qu'utilisateur, je souhaite... », permettent d'entrevoir clairement les attentes par rapport au module référentiel, de suivre leur évolution et de voir en quoi les technologies du web sémantique répondent à certains de ces besoins. Nous avons donc ajouté ces éléments à notre corpus de sources d'informations pertinentes pour l'étude de cas, en favorisant les aspects concernant l'intégration des vocabulaires SKOS, des notices d'autorités et des identifiants ARK.¹⁷⁹

II. L'approche qualitative pour compléter et approfondir les résultats

II.1. Le choix de l'entretien semi-directif

Un objectif : s'intéresser à la perception et au ressenti des archivistes et dépasser la vision « objective » des procédures

L'étude de cas, si elle permet effectivement de retracer l'historique du projet et d'en comprendre les différents enjeux, en analysant les différents choix qui ont été faits, n'offre pas réellement la possibilité de dépasser le stade de l'intentionnalité qui sous-tend l'élaboration du système d'archivage électronique. Puisque nous souhaitons aller au-delà de cette analyse, qui pourrait être qualifiée d'objective, des procédures du projet, nous avons estimé qu'une technique d'enquête de type qualitatif serait particulièrement adaptée, les entretiens permettant de traduire de manière beaucoup plus efficace et exhaustive les difficultés rencontrées ainsi que le point de vue et parfois même le ressenti des différents acteurs concernés. Par ailleurs, une approche quantitative, qui aurait pu être appliquée à l'approche comparative à laquelle nous avons initialement pensé, n'aurait pas ici permis de compléter notre analyse de manière satisfaisante. Dès lors que nous avons défini ce besoin d'exhaustivité et de précision, nous avons décidé de nous orienter vers des entretiens semi-directifs, puisque nous souhaitons donner l'opportunité à nos différents interlocuteurs de laisser libre cours à leur interprétation des différentes problématiques envisagées, tout en nous laissant la possibilité d'aborder les différents thèmes de manière spontanée, en fonction des propos tenus par les personnes interrogées, et de réagir à ces derniers.

Cependant, une telle méthodologie implique de prendre des précautions en matière

179 Archival Resource Key.

d'analyse et de restitution des résultats obtenus, puisque les propos tenus par nos interlocuteurs ne sont évidemment que le reflet de leur conception des différentes thématiques que nous souhaitions étudier. Il nous appartenait alors d'établir les liens entre ces différentes conceptions et de croiser les résultats obtenus afin de faire émerger une vision d'ensemble.

II.2. La réalisation du guide d'entretien

Le choix des questions et l'articulation avec les hypothèses

Parallèlement à la réflexion concernant le choix de nos interlocuteurs potentiels, il a été nécessaire de concevoir la trame d'un questionnaire, ou plutôt d'une grille d'entretien, qui se devait d'être le fil conducteur des entretiens. Celle-ci, puisque nous avons estimé qu'une approche semi-directive était préférable, devait être suffisamment ouvert pour pouvoir donner l'occasion aux professionnels interrogés d'étayer leur point de vue et nous permettre de creuser les points les plus pertinents. Dans cette optique, les différentes questions, que nous avons clairement formulées afin de pouvoir s'y référer aisément lors des entretiens tout en veillant à ne pas reprendre nos hypothèses mot pour mot, devaient permettre de mettre en relation les réponses obtenues avec les grands axes de notre travail. En ce qui concerne la formulation des questions, nous avons essayé, dans la mesure du possible, d'éviter les questions fermées, en préférant le recours à des formules interrogatives de type « dans quelle mesure » ?, « comment ? », etc. En effet, il ne s'agissait pas de confirmer ou d'infirmer certaines de nos hypothèses avec des réponses limitées à un simple « oui » ou à un « non », mais plutôt de laisser nos interlocuteurs développer leur perception des thèmes abordés.

Dès les prémices de notre réflexion concernant l'élaboration de notre grille d'entretien, nous avons déterminé les différents axes qui devaient y être représentés afin de jalonner l'entretien tout en permettant à nos interlocuteurs d'exprimer le fond de leur pensée. Ainsi, afin de contextualiser l'échange, de confirmer ce que nous pensions connaître de nos interlocuteurs et d'alimenter notre réflexion concernant la place de la culture professionnelle et des formations suivies sur la perception des évolutions de l'archivage et du web, nous avons dédié la première partie de notre grille d'entretien à

des questions relatives au parcours professionnel et à la formation suivie par les personnes interrogées. Des questions visant à positionner notre interlocuteur dans le projet du SAEM girondin et à établir ses interactions avec les autres membres et instances (décisionnelles et opérationnelles) du projet se devaient également d'y être intégrées, dans le but de poser des balises temporelles et de dresser le parallèle avec l'étude des procédures. Ensuite, il nous a fallu intégrer des questions relatives aux évolutions de l'archivage électronique et à leur impact sur les pratiques des archivistes, avant de dédier une partie de la grille d'entretien aux éléments relatifs aux interactions interprofessionnelles et à l'émergence de la gouvernance de l'information dans le contexte de l'administration électronique. Enfin, deux thèmes de notre trame d'entretien avaient pour objet l'articulation des technologies du web sémantique avec l'archivage, avec d'une part des interrogations concernant la compatibilité avec les pratiques actuelles des archivistes, et de l'autre des questions liées à leur potentiel pour les usagers.

Les différentes versions de cette grille d'entretien sont disponibles, dans leur intégralité, en annexe.¹⁸⁰

Un outil évolutif et prenant en compte de la pluralité des types d'acteurs interrogés

En raison de la diversité des personnes interrogées, nous avons estimé qu'une seule et unique grille d'entretien ne permettrait pas de couvrir l'ensemble des éléments que nous souhaitons recueillir, certaines notions n'étant que difficilement abordables par les prestataires de Logilab, par exemple, telles que les évolutions de l'archivage et l'interopérabilité avec les pratiques traditionnelles des archivistes. Réciproquement, du moins c'est ce que nous estimions les archivistes, n'étaient pas nécessairement les plus à même de nous apporter suffisamment de réponses concernant la conception du modèle de données et l'articulation des différentes couches du « layer cake » du web des données dans le cadre du projet. Pour autant, dans la perspective d'une analyse efficace des résultats obtenus, afin de garantir leur exploitabilité de garantir la cohérence de notre protocole, nous avons veillé à conserver une trame similaire, permettant de croiser les réponses de nos interlocuteurs et de tisser le lien entre leurs

180 Cf. annexes 6a et 6b.

visions et leurs approches des phénomènes étudiés.

Il a également été nécessaire de faire évoluer, sinon la trame d'entretien elle-même, notre manière de poser les questions et d'aborder les thèmes choisis. En effet, nous nous sommes rapidement rendu compte que les aspects technologiques, notamment en ce qui concerne les questions relatives au web sémantique, ne constituaient pas la meilleure porte d'entrée pour obtenir des réponses concluantes. Ainsi, il a fallu affiner nos questions et faire évoluer notre manière de procéder au fur et à mesure des entretiens. Dans cette optique, nous avons privilégié le recours à des exemples concrets, tels que les évolutions liées aux instruments de recherche ainsi qu'aux interfaces de consultation des sites web de services d'archives, mais aussi le mouvement des Humanités numériques, afin de cibler précisément la manière dont ces technologies peuvent jouer un rôle sur les pratiques des archivistes.

II.3. Définition du panel et déroulement des entretiens

Le choix des personnes interrogées : critères de sélection et objectifs ciblés

Dans la mesure où nous souhaitons avant tout consolider les éléments observés à partir de l'analyse des procédures du SAEM girondin, certains critères de sélection des répondants ont dû être privilégiés. En effet, nous avons recentré cette phase d'enquête qualitative sur un ensemble de professionnels qui sont directement intégrés au projet ou qui, en tant que prestataires et archivistes, ont pu intervenir de manière ponctuelle, et gravitent autour du dispositif. Nous avons donc ciblé, dans un premier temps, des acteurs qui avaient une expérience significative avec le projet et que nous estimions être en mesure de répondre à nos interrogations tout en nous permettant d'élargir nos observations afin de ne pas les circonscrire, de manière absolue, au projet du SAEM girondin. De ce fait, nous avons souhaité nous entretenir avec plusieurs archivistes, qui constituent le cœur de cible de notre étude :

- Agnès Vatican, directrice des Archives départementales de la Gironde, joue un rôle décisionnel au sein des COTECH (comité technique) et COPIL (comité de pilotage) du projet.
- Pascal Geneste, responsable du service collecte et traitement des Archives départementales de la Gironde, lui aussi membre des COTECH et COPIL.

- Delphine Jamet, archiviste du service Collecte et traitement des Archives départementales et membre du COPROJ (comité projet) du projet de SAEM, a un positionnement de spécialiste des métadonnées et de l'archivage électronique au sein de son service.
- Amélie Antoine, archiviste occupant la fonction de chef de projet Classement et conservation, intégrée au projet pour les questions relatives à l'archivage des délibérations des assemblées du Département et actes administratifs.

En outre, puisqu'un projet de refonte du site des Archives départementales était en cours, celui-ci étant intimement lié aux développements qui ont lieu dans le cadre du projet SAEM, nous avons jugé intéressant de prévoir un entretien avec Damien Duzert, agent de la DSIN (Direction des Systèmes d'Information et du Numérique) du Conseil départemental et responsable applicatif pour les outils de gestion documentaire utilisés au sein du Conseil départemental de la Gironde, dont Arkhéïa pour les Archives.

Un entretien avec Pascal Romain, « product owner »¹⁸¹ du projet, pour employer la terminologie propre aux méthodes agiles, a également été envisagé, notamment puisqu'il est un ancien archiviste qui travaille désormais auprès de la DSIN du Département de la Gironde. Ceci étant, puisque c'est lui qui encadrerait notre stage au sein du Département et que nous avons partagé, de manière détaillée, les objectifs et les hypothèses de notre travail avec lui, il a semblé difficile de passer outre le biais qu'introduit une telle relation et donc d'obtenir des résultats réellement exploitables.

Par ailleurs, et en raison de leur degré d'implication dans la conception de systèmes d'archivage électronique et puisqu'ils jouent un rôle de superviseur et subventionnent de nombreux projets de type SAE, nous avons envisagé de prendre contact avec les professionnels du SIAF avec lesquels l'équipe du SAEM a pu travailler, notamment dans le cadre de la mise en œuvre des profils SEDA et des notices producteur. Cependant, une telle démarche posait plusieurs problèmes, présentant notamment le risque de rester sur une approche conceptuelle des éléments que nous souhaitions analyser, et de ne constituer alors qu'un apport limité, dans la mesure où les éléments théoriques

¹⁸¹ Dans la méthodologie agile, le product owner est le responsable en charge de l'évolution d'une application ou d'un logiciel.

ont pu être abordé en première partie de ce travail.

En revanche, il nous a paru intéressant, pour ne pas dire indispensable, d'élargir le panel des personnes interrogées à d'autres typologies de professionnels, notamment aux éditeurs de logiciels et de solutions reposant sur les technologies du web sémantique qui ont pu intervenir dans le cadre du projet. Ainsi, une prise de contact avec les gérants de la société Anaphore, prestataire pour les services d'archives et éditeur du logiciel métier Arkheïa, utilisé par les Archives départementales de la Gironde, et de Logilab, le prestataire retenu pour l'élaboration du référentiel, que nous avons évoqué précédemment, a rapidement été envisagée. Nous avons alors souhaité interroger les personnes suivantes :

- Louis Colombani et Myriam Pauillac, dirigeants de la société Anaphore, ayant publié de nombreux billets de blog concernant l'évolution des instruments de recherche et les technologies du web sémantique.
- Sylvain Thénault, directeur de l'agence Midi-Pyrénées de la société Logilab et principal interlocuteur, du côté des prestataires, pour la conception et la réalisation technique du module référentiel du SAEM.

Malheureusement, nous n'avons pas pu obtenir de rendez-vous avec Louis Colombani et Myriam Pauillac de la société Anaphore, ceux-ci n'étant pas disponibles lors de la période dédiée à la conduite des entretiens. Ces entretiens auraient pu être enrichissants, dans la mesure où ils nous auraient permis d'aborder les questions relatives à la conception d'un modèle de données adapté aux besoins de l'archivage, puisque Anaphore a développé son propre modèle de données¹⁸², et d'analyser plus en détail l'interfaçage du module référentiel avec le logiciel métier des Archives départementales.

Prise de contact et modalités des entretiens

La prise de contact avec les professionnels travaillant au sein du Conseil départemental et des Archives départementales a été aisée et rapide et le fait d'être intégré directement au projet a permis à nos interlocuteurs d'avoir une compréhension

182 MDFA – Modèle de Description de Fonds d'Archives. Cf. <http://mdfa.anaphore.eu/0.8-content/mdfa.html>

facilitée de nos attentes par rapport à l'entretien. Ainsi, une simple demande de vive voix et la prise d'un rendez-vous via l'outil de calendrier utilisé par les services a permis de poser rapidement des rendez-vous. En ce qui concerne les différents prestataires et éditeurs de logiciels, nous avons choisi de formaliser la demande par le biais d'un e-mail reprenant les principaux objectifs du mémoire et précisant les modalités des entretiens. De ce fait, les entretiens menés auprès du premier groupe d'interlocuteurs ont pu être effectués de vive voix, dans les locaux du Conseil départemental et des Archives départementales, alors que l'entretien avec Sylvain Thénault a été mené en visioconférence, afin de pouvoir enregistrer et retranscrire ultérieurement l'ensemble de l'échange.

Les premières minutes de l'entretien, avant même de procéder à l'enregistrement de l'échange, étaient consacrées à la présentation des objectifs de notre travail et des différents thèmes abordés dans la grille d'entretien. Ce fut alors l'occasion de préciser que l'entretien allait être enregistré, de demander évidemment l'accord de notre interlocuteur, et de rappeler que les données recueillies pouvaient être anonymisées si la personne interrogée le souhaitait.

Les entretiens, conformément à nos estimations initiales, ont duré en moyenne environ quarante-cinq minutes, avec plusieurs d'entre eux, auprès de certains archivistes, qui ont pu durer près d'une heure complète. Nous avons également tâché, dans la mesure du possible et tant que les réponses ne s'écartaient pas de la trame prédéfinie, de ne pas interrompre nos interlocuteurs et de les laisser aller jusqu'au bout de leur pensée. Pour chacun des entretiens, nous avons également mené des recherches concernant le profil de notre interlocuteur et préparé des ressources, notamment en nous référant à des documents ou exemples de référence tels que la feuille de route du ministère de la Culture sur les métadonnées culturelles, le rapport Nougaret¹⁸³ ou encore le portail FranceArchives, permettant d'approfondir certains points au moment de l'échange.

183 NOUGARET, Christine, [2017]. Une stratégie nationale pour la collecte et l'accès aux archives à l'ère numérique. In : FranceArchives [en ligne]. [Consulté le 10 août 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/28204701>.

III. Procédure d'analyse et d'interprétation des résultats

La mise en regard des données issues de l'étude des procédures avec les résultats des entretiens

Étant donné que nous souhaitons enrichir notre étude de cas à l'aide des propos recueillis lors des entretiens, nous avons estimé qu'il était préférable de traiter simultanément les résultats de nos analyses, c'est-à-dire de ne pas séparer arbitrairement ces éléments mais plutôt de les coupler afin d'être en mesure de confirmer ou d'infirmer nos hypothèses.

Cependant, certains aspects n'ont réellement été analysables qu'à partir des témoignages des professionnels interrogés, puisque relatifs à des thèmes qui vont au-delà de la simple mise en œuvre technique d'un système d'archivage électronique. De ce fait, en ce qui concerne de nombreux points que nous souhaitons aborder, notamment les enjeux liés aux interactions interprofessionnelles et la perception des évolutions de la pratique de l'archivage, les éléments issus des entretiens constituent la principale source de données et ont pu alors prendre le pas sur les résultats de notre analyse des procédures du projet.

En ce qui concerne les entretiens, nous avons opté pour une transcription intégrale à partir des enregistrements, précédée d'une prise de notes au moment même des entretiens, afin de souligner les temps forts de l'entretien et les points les plus intéressants afin d'en faciliter l'analyse par la suite. Il s'agissait également d'être en mesure, en prenant les précautions nécessaires, de pouvoir interpréter les non-dits et plus rarement certaines hésitations et répétitions qui peuvent mettre l'accent sur des éléments cruciaux. La transcription intégrale des différents entretiens est par ailleurs disponible en annexe à ce mémoire.¹⁸⁴

Par ailleurs, nous avons fait le choix, de la même manière que lors de la conception de notre grille d'entretien, de rattacher l'ensemble des résultats recueillis, qu'il s'agisse des éléments factuels issus de l'étude de cas ou des données extraites à partir des entretiens, à des thématiques qui correspondent aux grands axes de notre travail, et donc aux hypothèses et aux idées que nous émettions en première partie.

184 Cf. annexe 7a à 7f.

Troisième partie : résultats de l'expérimentation

Les données issues du traitement de la documentation et des ressources accumulées lors du développement du SAEM girondin, ainsi que celles que nous avons pu récupérer à partir des différents entretiens, nous ont permis de construire un corpus conséquent, dont l'objectif est de pouvoir confronter nos postulats de départ avec les réalités observées sur le terrain et les témoignages des professionnels. Pour ce faire, nous avons choisi d'articuler ces différents éléments en fonction des principales thématiques que nous avons identifiées préalablement, tout en laissant place aux aspects plus inattendus et aux questions subsidiaires qui ont pu émerger au fil de notre expérimentation. Dans une première sous-partie, nous analyserons l'impact du développement de l'archivage électronique sur les pratiques des archivistes en laissant notamment la part belle aux points de vue des professionnels en question. Ensuite, c'est au développement d'une gouvernance transversale de l'information que nous intéresserons, en établissant le lien entre les objectifs et la structuration du projet SAEM avec le ressenti des acteurs impliqués. Enfin, nous nous attacherons à démontrer la manière dont les technologies du web sémantique s'inscrivent dans ces mutations et à déterminer comment elles peuvent être intégrées à un projet d'archivage électronique, avant d'étudier la façon dont leur potentiel et leurs implications technologiques sont perçus par les professionnels interrogés.

I. La transition numérique des services d'archives, symbole d'une bascule dans les pratiques professionnelles

I.1. L'archivage électronique, signe avant-coureur d'une montée en compétences inéluctable ?

Nous supposons, à l'issue de notre première partie, que les bouleversements liés à l'essor du numérique et du développement d'outils reposant sur des schémas de métadonnées normalisés, couplés à des impératifs de plus en plus stricts concernant les documents nativement numériques, constitueraient le point de départ d'un certain

nombre de mutations dans les pratiques des archivistes. Les tableaux de travail mis en place par l'équipe du SAEM concernant les procédures de pré-versement ainsi que le manuel utilisateur de l'outil, rappellent la volonté de faire coïncider la méthodologie employée dans le cadre du traitement des archives numériques avec des pratiques archivistiques bien implantées. Ainsi, on y retrouve une succession d'opérations qui répondent au fonctionnement habituel du traitement d'un fonds d'archives, avec quelques spécificités et variations qui caractérisent les impératifs supplémentaires liés à l'archivage de données et de documents numériques. Parmi ces opérations sont notamment évoquées la « préparation en amont, [le] travail avec le producteur », le fait d'« évaluer, sélectionner », de « capter les données » et de « traiter les données en vue du transfert. » De la même manière, lors de notre échange avec Amélie Antoine, celle-ci évoque son rôle de la manière suivante : « pour moi, [mon rôle] c'est de définir... dans le jargon classique, on dirait le plan de classement. Mais mon but c'est d'identifier les métadonnées à récupérer. » Il y a donc une mise en relation évidente entre les pratiques traditionnelles des archivistes et celles qui peuvent émerger avec le traitement de documents d'archives numériques.

Pareillement, le guide utilisateur rappelle que « fréquemment, la production administrative est regroupée au sein de dossiers, car on utilise par analogie l'unité intellectuelle que l'on traite et le support au sein de laquelle on la classe » et signale que le fonctionnement du référentiel repose sur l'intégration des unités d'archives au sein de profils d'archivage qui permettent de déterminer les règles applicables à ces dossiers. De la même manière, la théorie des trois âges des archives trouve également un écho à travers l'articulation des différentes briques du dispositif girondin, puisque chaque couche applicative peut être mise en correspondance avec une étape du cycle de vie d'un document d'archives.¹⁸⁵

Par ailleurs, Delphine Jamet, archiviste du service collecte et traitement, rappelle que « la grosse mutation s'est faite avec la mise en ligne des documents numérisés, déjà, et la rétro-conversion des instruments de recherche. » Cela laisse supposer que si le développement de l'archivage électronique implique de nouveaux procédés et de nouveaux outils et s'il demeure un prisme à travers lequel il est possible d'observer les

185 Cf. schéma en annexe 12.

mutations liées au numérique, il ne constitue pas le point de départ de ces dernières dans le monde des archives.

En outre, Pascal Geneste signale que les documents numérisés ou archivés électroniquement ne constituent qu'« une goutte d'eau par rapport à la production extrêmement massive de l'administration ». Il s'agit là d'un facteur permettant, entre autres, d'expliquer pourquoi les professionnels des archives semblent avoir des difficultés à percevoir, encore aujourd'hui, l'impact de l'archivage électronique et de l'accès aux ressources en version numérique, puisque les documents numérisés ou archivés électroniquement ne constituent qu'une part très faible de la production documentaire des administrations. De plus, ce sont uniquement certains types de documents qui sont directement concernés par ces procédures, comme le rappelle Delphine Jamet : « Pour l'instant, l'archivage électronique, ce n'est pas forcément des documents d'histoire, enfin, c'est vraiment de la production courante, enfin récente ». Dans le cadre du SAEM girondin, le nombre de flux restreint (au nombre de quatre au moment de la mise en production), au regard de la volumétrie de l'ensemble des documents à traiter, contribue également à ce sentiment d'un périmètre limité de l'archivage électronique et, par extension, au manque de visibilité que nous évoquions. Le fait que les ambitions et les attentes par rapport à l'archivage électronique ne soient pas récentes¹⁸⁶, sans pour autant avoir réellement mené à des systèmes concrets et robustes, et le fait que la diversité des documents à archiver n'est souvent que partiellement prise en compte, renforcent ce sentiment. C'est en tout cas ce qu'affirme P. Geneste :

« Ces quatre flux ne représentent évidemment pas l'ensemble de la production électronique que peut produire l'administration [...] donc finalement, on ne travaille que sur quatre flux, des 10 % de la collecte traditionnelle. »

Dans les faits, au-delà des questions liées au développement de ces systèmes, c'est l'ensemble des processus propres au fonctionnement du numérique qui est à l'origine de mutations dans les services d'archives, comme partout ailleurs. En effet, l'explosion de la volumétrie documentaire, avec des documents qui peuvent être dupliqués à

186 « À la fin des années 90, début des années 2000, on parlait déjà d'archivage électronique et de développer des systèmes d'archivage électronique, qui n'ont jamais abouti [...] Il va bien falloir élargir une fois que les premiers systèmes seront en place, élargir ce champ, c'est pour ça que je parlais d'une trentaine ou quarantaine d'années, pour progressivement arriver à absorber la totalité de la production électronique. »

l'infini, est une source de complexité supplémentaire. D'après Pascal Geneste, c'est la notion même de patrimoine, et par extension celle d'archives, qui ont été considérablement élargies avec l'essor du numérique :

« On a assisté à un élargissement considérable de la notion de patrimoine [...] et à l'intérieur de cette notion-là, il y a les archives évidemment mais aujourd'hui tout est archive. » « [Il s'agit de] réduire les typologies, qui ont explosé. C'est vrai pour le papier et peut être encore plus pour l'électronique, puisqu'on est là avec un outil qui nous permet de multiplier la production, par le biais simple du copier-coller, mais aussi parce qu'à chaque fois qu'on produit une version supplémentaire d'un document, on n'écrase pas la précédente, on l'ajoute en quelque sorte à la précédente. »

Les propos tenus par notre interlocuteur laissent alors supposer qu'une rationalisation et une simplification de la production archivistique semble alors nécessaire. Suite à l'élargissement massif de la notion d'archives, il semblerait désormais préférable d'en venir à une réduction des typologies documentaires et à une révision du champ d'action des services d'archives :

« On risque d'assister dans les années à venir à une réduction de cette notion-là, à une réduction aussi certainement des typologies documentaires dont je parlais tout à l'heure, et à une réduction du champ d'action des services d'archives, qui ne pourront pas absorber la totalité de ce qui est produit. »

Ce sont également certains outils, encore au cœur des pratiques des archivistes, qui sont remis en question, notamment en ce qui concerne certains cadres de classement et référentiels. C'est en effet la pertinence et l'actualité de ces documents utilisés par les archivistes pour classer et traiter les fonds qui sont remises en cause. Pascal Geneste s'interroge alors : « On présente les données suivant un cadre de classement des archives départementales qui remonte quand même à 1840. Donc, en 2020, est-ce qu'il est toujours pertinent de présenter, notamment les fonds que nous conservons, suivant un cadre de classement qui remonte presque à 180 ans ? »

Cette remise en cause rejoint le questionnement de l'équipe du SIAF, lors d'une réunion avec le comité projet du SAEM girondin, concernant le recours à la série W, dite continue, pour les archives numériques :

« Les environnements de l'archivage papier et de l'archivage numérique étant très différents, un certain nombre de pratiques professionnelles sont difficilement transposables. C'est en particulier le cas de la cotation. La série W, réglementaire pour les archives contemporaines, est-elle opérante pour les documents électroniques ? Si tel est le cas, selon quelles modalités ? Dans l'hypothèse inverse, quel identifiant utiliser tant pour l'archivage intermédiaire que pour l'archivage

définitif dans le domaine numérique ? »

Comme nous l'avons évoqué dans notre première partie, c'est également le régime de collecte des archives qui se voit bouleversé par le numérique. En évoquant le rapport Nougaret, Agnès Vatican évoque cette rupture de la manière suivante : « Sur le numérique, il pose un certain nombre de questions, effectivement des problématiques de rupture entre le producteur et le territoire, la notion d'original, la question de la sélection qui se pose plus que jamais, effectivement, face au volume de données nativement numériques. » En effet, les archives numériques, et c'est d'autant plus vrai avec l'explosion des documents nativement numériques, complexifient le dialogue entre services producteurs et services d'archives et impliquent de nouvelles modalités de sélection.

Une incompatibilité entre le rythme des évolutions technologiques et le besoin de stabilité des services d'archives ?

Le rythme des évolutions qu'imposent la transition numérique et l'essor de l'archivage électronique constitue également un facteur d'inquiétude, dans la mesure où les agents doivent être formés et se familiariser à des concepts qui peuvent déjà être perçus comme obsolètes au moment de leur arrivée dans les services. C'est en effet ce qu'explique Agnès Vatican : « Je pense qu'il faut arrêter de penser que les professionnels peuvent trouver, en plus dans un contexte où les ressources en moyens sont comptées, le temps de se former, tous les 10 ans, à un nouveau modèle de données. » Par ailleurs, cette difficulté est là aussi accentuée par les problématiques liées à la volumétrie et aux nombreux chantiers, en termes de traitement électronique des fonds papier, qui ne sont pas encore clos malgré un recours à des technologies qui sont en place depuis plus de 15 ans : « On a encore des fonds qui ne sont pas classés, qui n'ont pas d'instrument de recherche, donc se dire : tout ce qui était en EAD, on va le repasser dans un nouveau modèle de description etc, on va tout reprendre, alors qu'on a des fonds par ailleurs qui n'ont pas d'instruments de recherche. » Cette incompatibilité se retrouve, là aussi d'après Agnès Vatican, également au niveau même du fonctionnement des services d'archives et plus largement des administrations, puisqu'elle souligne les besoins de stabilité et de pérennité qui permettent le fonctionnement de ces services :

« C'est quand même un peu une course perpétuelle, on est des institutions qui nous inscrivent dans une certaine pérennité, on a deux siècles derrière nous de fonctionnement [...] Je pense qu'aujourd'hui, où l'administration est de plus en plus sur une recherche de performance, et qui dit performance dit outils adaptés et aussi dans un certain nombre de cas continuité dans le fonctionnement. »

Pour autant, malgré la nécessité de prendre en main de nouveaux outils et d'acquérir un vocabulaire technique riche, les archivistes interrogés ne semblent guère inquiétés par la transposition des modalités de l'archivage numérique aux pratiques papier. Il convient alors de nous intéresser aux facteurs qui limitent ces difficultés.

La normalisation, garante de l'interopérabilité entre les pratiques de l'archivage papier et les apports du numérique

Nous avons mis en lumière, dans notre première partie, l'importance du recours à des schémas de métadonnées interopérables avec les normes de description archivistique préexistantes. Pour autant, nous ne soupçonnions par nécessairement alors la manière dont ces nouveaux outils allaient permettre de formater et de structurer les pratiques des archivistes. Plutôt qu'une source de difficulté, cette nécessité d'intégrer la description archivistique à l'environnement numérique est désormais perçue comme le point de départ d'une uniformisation des processus métier des archivistes. Ainsi, Pascal Geneste résume cette évolution de la manière suivante :

« Cet environnement numérique, me semble-t-il, a permis une structuration de la description archivistique que l'on n'avait pas forcément auparavant, que l'on n'appliquait pas de manière aussi systématique auparavant, [...] l'évolution technologique à laquelle on assiste nous contraint, et contraint les archivistes, à utiliser à la fois beaucoup plus les normes et la structuration de la description qu'on ne le faisait auparavant. »

Cette nécessité ne répond pas uniquement aux impératifs des logiciels métier, dont le fonctionnement et l'interface dépendent souvent de ces normes de structuration des données, mais découle également des besoins d'accès aux fonds à travers les sites et portails d'archives : « On est obligé de normaliser et de structurer sa description et ses données, sinon vous ne passez pas le portail en quelque sorte, vous restez à la porte. » C'est un point de vue qui est partagé par l'ensemble des archivistes que nous avons interrogés, et d'après Agnès Vatican, ces observations dépassent le cadre de la description archivistique, l'informatisation des services, dans sa globalité, ayant permis d'imposer un fonctionnement « plus rigoureux » en ce qui concerne les tâches de

gestion courante.¹⁸⁷

Dans le même temps, le recours à des schémas de métadonnées standardisés et à des normes spécifiques a également permis de préparer le terrain et de poser les bases d'une culture de la métadonnée dans les services d'archives : « C'est plus ou moins évident, je pense que celle qui a le plus facilement, enfin facilement ce n'est pas vraiment le mot, mais qui est peut-être le plus entré en vigueur, c'est ISAD(G) et donc l'EAD. » Ainsi, lorsque nous interrogeons Agnès Vatican au sujet de la transition vers le modèle Records In context conçu par l'ICA, celle-ci estime que « le fait d'être passé à l'EAD peut préparer aux étapes suivantes », soulignant ainsi qu'une culture de la description « fragmentée » et de la métadonnée serait déjà en construction dans les services. Il est également intéressant de noter que ce contexte de normalisation des processus de description archivistique, imposé par le numérique, a eu un impact direct sur les pratiques concernant les documents papier, comme l'explique Delphine Jamet :

« Avec le fait qu'on aille sur de l'archivage électronique, on a été obligé de mieux contextualiser la production documentaire et donc du coup ça a impacté nos pratiques sur le papier. Parce que si on le fait pour le numérique, alors pourquoi ne pas le faire pour le papier, donc on a mis en œuvre des procédures pour pouvoir faire notamment les notices d'autorité. »

Au sujet des notices d'autorité producteurs, dont l'intégration au format EAC-CPF est l'une des composantes cruciales de l'harmonisation via le module référentiel, elle ajoute : « De façon normalisée, on ne les faisait pas avant, donc c'est un vrai changement. » En effet, si les pratiques papier influencent naturellement la manière d'appréhender le document numérique, les nouveaux aspects qui doivent être pris en compte pour ce dernier, notamment en ce qui concerne sa provenance et sa contextualisation, ont rétroactivement un impact sur les pratiques plus traditionnelles. De fait, c'est l'exemple des notices d'autorité, qui relèvent d'une pratique qui découle de l'archivage électronique et de la norme ISAAR-CPF (et du schéma de métadonnées EAC), qui est ici évoqué pour illustrer ces évolutions.

L'uniformisation des processus engendrée par le développement de la normalisation

187 « [Avec le numérique et la mise en ligne de documents d'archives] il y a eu évidemment davantage de formalisme, davantage de normalisation, ça c'est inhérent, mais dès qu'il y a eu informatisation, même sur les aspects de gestion, sans parler de la description documentaire, quand vous avez une gestion de salle de lecture informatisée, il y a forcément quelque chose qui se met en place, qui est plus rigoureux, qui est plus normé »

démontre également l'importance des initiatives aux niveaux central et international. D'après Pascal Geneste, ce sont ces impulsions menées au niveau du SIAF et par la communauté internationale des archivistes qui ont permis ce changement de paradigme : « Les opérations de coordination qui sont développées au niveau central et qui ne l'étaient peut-être pas, ou moins, auparavant, [...] permettent de s'inscrire dans cette normalisation et structuration des données. »

En outre, la politique d'archivage du SAEM girondin rappelle que « la production d'une archive, son versement et sa gestion s'organisent autour d'une chaîne d'intervenants ». C'est le modèle OAIS, que nous avons évoqué en première partie, qui précise le rôle et le positionnement de chacun de ces intervenants. Chacun des modules du SAEM et les actions qui y sont associées trouvent alors leur place au sein des différentes entités du modèle OAIS.¹⁸⁸ Les équipes du dispositif girondin ont alors, notamment pour encadrer la mutualisation des processus, décidé d'établir une mise en correspondance entre les différentes entités impliquées dans le projet SAEM avec les différents rôles décrits par le modèle OAIS.¹⁸⁹ Nous le voyons, les modèles conceptuels et standards de l'archivage électronique, notamment le modèle OAIS sur lequel l'architecture du SAEM girondin repose entièrement, ont été conçus dans l'optique d'une interopérabilité avec les étapes du traitement des documents d'archives papier, afin de prendre en compte les procédures déjà en œuvre dans les services d'archives d'horizons divers :

« Quand on regarde en tout cas le modèle OAIS, finalement c'est la chaîne archivistique, avec la planification, la pérennisation et la préservation. Pour moi, ça rentre complètement, quand on regarde le schéma, c'est ce qu'on doit faire aussi pour le papier, [...] on a tout ce qu'il faut [pour être] en conformité avec le modèle OAIS. Donc ça n'a pas vraiment révolutionné tout ça. »

La « chaîne archivistique », comme le signale ici Delphine Jamet, est alors respectée et il y ne semble pas y avoir de dissonance entre la manière dont est pensée le cycle de vie d'un document d'archive au sein d'un service public d'archives et le positionnement des différentes briques du modèle OAIS. D'après Amélie Antoine, les difficultés en matière de traitement des archives numériques se retrouvent davantage au niveau de la conservation de ces documents, puisque ceux-ci doivent garder leur valeur probante et, tout en demeurant lisibles et exploitables par les outils de demain : « Pour moi, il y

188 Cf. schéma en annexe 13.

189 Cf. répartition des rôles pour la mise en production du SAEM girondin en annexe 14.

a toujours des objets, alors que ce soit sur un support électronique... En fait, ce qui change, c'est la question de la conservation. »

D'après les archivistes qui ont répondu à nos questions, l'archivage électronique ne représenterait alors qu'un bouleversement très limité en termes de processus et de conception de l'objet archive, comme le rappelle là aussi Amélie Antoine :

« Les archives électroniques, fondamentalement, ne changent pas le métier de l'archiviste. Fondamentalement, non, sur des aspects techniques, technologiques, oui [...] clairement, entre le papier et l'électronique, le sens du métier est le même. »
« C'est une transposition. Comme on le disait tout à l'heure, le vocabulaire évolue, certes, il faut se l'approprier et que nos cerveaux... mais oui, pour moi il n'y a aucun problème de corrélation, [...] le système n'est pas révolutionnaire en soi... »

Pour autant, ce même témoignage met clairement en lumière un besoin d'appropriation d'un vocabulaire spécifique et une évolution considérable sur le plan technologique, qui joue, in fine, un rôle considérable sur le plan des compétences que les archivistes doivent développer : « Il y a des adaptations, en termes de compétences à acquérir, [...] pour beaucoup de collègues, c'est une restructuration très importante. » Delphine Jamet va quant à elle plus loin, en rappelant qu'avant ce processus de normalisation, l'archivage s'effectuait de manière « instinctive », avec une réflexion qui pouvait être propre à un bureau ou à un service. Il s'agit alors pour les archivistes de « s'approprier le langage, le vocabulaire, qui est plus complexe ». Dans les faits, ces différents standards et modèles, qui ont tendance à se multiplier, imposent aux archivistes d'acquérir un socle, en termes de terminologie, plus rigoureux et précis.

D'après ces différents témoignages, la transition vers l'archivage électronique et la gestion des métadonnées passe avant tout par le plan technologique, mais ne semble pas constituer pas une réelle évolution en termes de pratiques professionnelles. La volonté de transposition des normes liées à l'archivage électronique aux pratiques déjà mises en œuvre pour la gestion d'archives papier limite ce besoin d'adaptation. Ceci étant, les archivistes, notamment ceux qui sont moins à l'aise avec l'outil informatique et les enjeux de l'administration électronique, doivent intégrer une nouvelle terminologie et peuvent percevoir cette évolution comme un bouleversement de leurs pratiques.

I.2. Des besoins en termes d'acculturation et de conduite du changement

L'appropriation des outils de l'archivage électronique par les archivistes : l'enjeu de la conduite du changement

Ayant débuté en 2011, le projet du SAEM girondin ne fera vraisemblablement l'objet d'une mise en production qu'à partir de la fin d'année 2017¹⁹⁰, et c'est alors la question de la nécessité d'une acculturation et d'un accompagnement au changement qui est mise en lumière par les personnes avec lesquelles nous nous sommes entretenues. Ainsi, Sylvain Thénault s'inquiète de l'appropriation tardive de ces outils par les utilisateurs finaux :

« Le vrai saut c'est quand on va passer à des archivistes plus grand public, qui ont moins l'approche technique que peuvent avoir tous les gens qui sont dans le COPROJ aujourd'hui. » « Pour moi, ça fait partie des... je ne sais pas si je dois dire échecs, mais en tout cas des problèmes potentiels [...] On a bien réussi l'aspect [...] humain entre les équipes de dev et les équipes projet mais par contre on a raté un truc, c'est que ça aurait dû être plus ou moins déployé, entre les mains des utilisateurs finaux, depuis longtemps »

Cette problématique de la représentativité des membres du COPROJ par rapport aux utilisateurs finaux pose alors la question de la capacité des archivistes à manipuler ces outils et des dispositifs qui devront être mis en place pour effectuer cette transition.¹⁹¹ Agnès Vatican, directrice des Archives départementales de la Gironde, se pose également la question de l'appropriation des outils développés dans le cadre du SAEM, par les services auxquels ils se destinent, à savoir notamment les archivistes, qui pour l'instant n'ont pas eu l'occasion de manipuler l'outil : « La question qui se pose pour moi aujourd'hui, c'est plutôt comment on passe de ce mode projet, très intégré, très performant, à la mise en production et à l'acculturation des équipes, parce que derrière, il s'agit bien que les équipes s'approprient l'outil et l'utilisent ». D'après elle, « il y a encore un travail de conduite du changement à mener. » Le terme d'acculturation souligne ici que cette transition dépasse la simple acquisition de compétences techniques mais que ces aspects doivent être intégrés dans les pratiques quotidiennes des équipes. De ce fait, cette transition semble être une véritable source

190 Cf. planning du projet SAEM en annexe 17.

191 Toujours d'après S. Thénault : « C'est un problème, aujourd'hui, puisque les pseudo-utilisateurs finaux ne sont plus du tout représentatifs puisque c'est les membres de l'équipe projet qui sont surformés sur le sujet et sur les technologies sous-jacentes »

d'interrogation, dans la mesure où une relative incertitude plane encore quant à la capacité, en l'état actuel, des archivistes à s'approprier l'ensemble du dispositif. Par ailleurs, la mise en place d'un tel système, et nous pouvons supposer que cela peut-être le cas dans n'importe quel service qui doit apprendre à composer avec un nouveau SAE, peut avoir des conséquences importantes sur le fonctionnement quotidien des agents qui doivent l'utiliser, d'où la nécessité d'assurer la conduite du changement qui est fréquemment évoquée par nos différents interlocuteurs. Là aussi, Sylvain Thénault souligne cette problématique et s'interroge dans le même temps sur la portée effective des outils conçus :

« Ça devrait être à des gens qui, potentiellement, vont arriver dessus, ne vont pas trop savoir, vont vouloir l'utiliser, ils vont avoir des choses qu'ils faisaient au quotidien qui vont maintenant être faites par là ». « Aujourd'hui, je ne sais pas complètement dire ne serait-ce que qui est le public exactement derrière tout ça et qu'est-ce qu'attend ce public, [...] ça se destine aux archivistes, alors qu'au départ on pouvait voir des volets avec des utilisateurs plus grand public, par exemple. »

Ces observations peuvent être élargies à l'ensemble de l'archivage électronique et des facettes du métier d'archiviste qui impliquent une connaissance plus approfondie des technologies, comme le rappelle Pascal Geneste, qui insiste sur l'importance d'une montée en compétences des utilisateurs finaux, c'est-à-dire les agents des services d'archives concernés :

« Nous avons pour ambition, dans les mois et années à venir, de faire monter en compétence les nombreux agents. L'idée, c'est qu'il y ait une appropriation de beaucoup plus d'agents sur cette question-là, et sur le système qui aura été développé, et sur son utilisation, pour que progressivement cela infuse auprès des producteurs de documents. »

C'est la pérennité du système qui est en jeu à travers ce besoin de montée en compétences, puisque l'ensemble des agents doit être en mesure de les manipuler et de comprendre le fonctionnement afin que ceux-ci s'inscrivent véritablement dans les pratiques courantes des archivistes. Toujours selon Pascal Geneste, un tel accompagnement au changement n'est pas toujours facile à mettre en place.¹⁹²

192 « Derrière tout ça il y a quand même les agents, les fonctionnaires qui sont là, soit depuis peu soit depuis longtemps, et dont il faut prendre la tâche. Il y a l'accompagnement au changement, qui est une donnée importante et dont on parle, comme ça, en termes très théoriques mais, sur le plan pratique, n'est pas toujours mis en œuvre. »

Du document à la donnée : quelle place pour l'instrument de recherche ?

Comme nous l'évoquions en première partie, l'élaboration des instruments de recherche et la manière dont les archivistes conçoivent le document en tant que tel pourraient constituer un point de tension et un obstacle à la transition numérique. Ainsi, lorsque nous interrogeons Delphine Jamet par rapport aux mutations en termes de granularité de l'information à traiter par les archivistes, elle nous répond de la manière suivante : « Pour moi un document, que ça soit une pièce, un groupe de pièces, un dossier, c'est plutôt là où est la granularité pour moi ». Cette transition du document à la donnée semble alors difficilement perçue au niveau des instruments de recherche, la question de la granularité de l'information semble encore coïncider avec la manière dont les archivistes traitent les documents, avec une description à plusieurs niveaux. Par ailleurs, ses propos permettent de préciser cette vision, qui est spécifique aux archivistes, et de souligner les difficultés en termes de faisabilité : « On va être obligé d'aller à la pièce, enfin plus je veux dire, mais ça se fera automatiquement en fait, d'où l'importance d'être en amont de la chaîne », avant d'ajouter : « mais on ira pas pour tout à la pièce, [...], je pense qu'on va perdre beaucoup de choses, on ne fera pas l'histoire comme on a fait l'histoire jusqu'à présent ». Ainsi, le besoin en termes d'exhaustivité et de précision de la description qu'impliquent le numérique et la mise en ligne des données issues de l'archivage est bel et bien identifié. Il s'agit alors d'une difficulté supplémentaire, puisque la description ne se fait que rarement au niveau de la pièce, dans le domaine des archives, pour des raisons de faisabilité, comme nous l'avons évoqué en mentionnant la volumétrie des documents conservés.

De plus, l'instrument de recherche, notamment dans sa forme électronique, est quant à lui toujours considéré comme pertinent, en tant qu'outil de description des fonds. Pour autant, le fait qu'il ne faille pas le lire « comme un livre »¹⁹³ indique que sa valeur dépasse celle d'un simple document ou inventaire linéaire, mais que c'est la présence même des données qui s'y trouvent qui permettent l'accès à l'information. D'après nos observations et les propos recueillis auprès des archivistes interrogés, le document numérique ne bouleverse pas fondamentalement le traitement archivistique mais la

193 D'après A. Antoine : « L'instrument de recherche en archives, c'est quelque chose de très complet, surtout si on parle d'instrument de recherche électronique, qui s'en sert totalement dans cet esprit d'échange, [...] ce n'est pas comme un livre, l'instrument de recherche donne de l'information dès le début, dès le plan de classement. »

question du support suppose évidemment un changement dans la manière de percevoir le contenu du document. Amélie Antoine présente cette distinction de la manière suivante : « C'est vrai que quand tu penses papier, tu identifies directement, tu vas lire les informations dessus directement, tandis qu'un document numérique, la matérialité n'est pas la même. Non, la donnée en elle-même et les métadonnées ont du sens. Par contre, il faut arriver à les percevoir. »

Agnès Vatican, en revanche, se méfie d'une opposition radicale entre document et donnée, et considère, ce qui est peu surprenant au regard de la place qu'occupe le document dans la pratique de l'archivage, que la forme documentaire reste encore largement d'actualité : « Je pense que l'un ne doit pas exclure l'autre [...], La forme livre, elle reste opérante à bien des égards [...] parce qu'effectivement, quand on a en tête la référence de l'article, c'est tellement simple d'attraper le volume et de le feuilleter. » Cette conception du document, abordée à cette occasion à travers l'exemple de l'instrument de recherche, correspond aux habitudes de lecture de type « littéraire », linéaire, qui demeurent d'après notre interlocutrice largement d'actualité aujourd'hui pour les publics des archives. Il s'agirait alors « de trouver des articulations » entre ces manières de concevoir l'information et les dispositifs d'accès à celle-ci. Pour autant, toujours après elle, « penser que tout deviendra donnée, à mon avis c'est une utopie un peu dangereuse. » Pour résumer ces différents points de vue, la place du document, au niveau sa production et en tant qu'objet fruit de l'activité humaine, n'est pas remise en question, c'est plutôt la question de son accès et de l'exposition des informations qui s'y trouvent qui doit être réévaluée.

Des carences en termes de formation et de sensibilisation à la donnée et à l'archivage numérique ?

Les différentes formations suivies par les archivistes par rapport aux questions de l'archivage électronique et à la gestion de métadonnées, qu'il s'agisse des formations initiales ou des différents séminaires auxquels ils assistent, constituent un élément essentiel pour comprendre la manière dont est forgée leur perception de ces questions. Amélie Antoine, cheffe de projet conservation et classement, regrette ainsi la nature de ces formations, qui laissent bien souvent de côté les aspects techniques et les enjeux concrets, pour les archivistes mais aussi pour les publics :

« Jusqu'à la dernière [formation], d'ailleurs, j'étais déçue, puisque je faisais beaucoup de formation à chaque fois je ne suis pas la seule à montrer mon mécontentement parce que jusqu'à présent, on parlait beaucoup des normes, de lois, c'est très juridique tout ça mais ça manquait de concret et on se retrouvait souvent, avec son papier et son crayon, à suivre des Powerpoint. »

Ces propos soulignent également le fait que les archivistes ne sont que rarement amenés à manipuler des outils et à développer des compétences techniques au cours de ces formations. Cela laisse supposer que les archivistes ont besoin d'être mis face aux outils et de pratiquer afin d'avoir une vision concrète de ce qui est réalisable, et rejoint l'idée qu'un certain manque de visibilité règne encore sur les apports de l'archivage électronique pour les professionnels qui sont pourtant les plus à même de s'emparer de ces questions. Agnès Vatican ajoute par ailleurs que l'apprentissage de ces notions se fait avant tout par le biais d'une confrontation sur le terrain : « En formation initiale, il y avait assez peu de choses, on avait quand même à l'école des Chartes une formation aux langages informatiques [...] le fait d'avoir conduit deux informatisations, en quelque sorte, c'est quand même quelque chose qui permet d'apprendre pas mal. » Il convient toutefois de nuancer ces propos, puisque nous n'avons pas fait une étude des différentes formations en archivistique qui existent actuellement, et parce qu'ils reflètent davantage le point de vue des décideurs. Ceci étant, ces éléments rejoignent également le point de vue d'Amélie Antoine, qui insistait sur la nature des formations, très souvent portées sur les différents formats et supports plutôt que sur les finalités et les outils. Cela dit, ces difficultés ne sont pas uniquement le fait des formations suivies mais, comme nous le supposons, découlent parfois d'une culture professionnelle, caractérisée par un rapport particulier avec le document en tant que support d'information. Ainsi, Pascal Geneste met en avant un éventuel facteur générationnel pour expliquer ces difficultés :

« Je suis d'une génération qui n'a pas [...] grandi avec l'informatique [...] je suis d'une génération qui pense, de toute manière, qu'il faudra au moins une bonne trentaine d'années avant que cette mutation entre la production traditionnelle, la production papier, qui est une tradition très ancienne, ne soit pleinement prise en compte »

Si nous ne pouvons vérifier la validité de ce facteur générationnel à partir de notre protocole d'expérimentation, il s'agit tout de même d'interroger l'influence d'un tel aspect sur la place qu'occupent les considérations liées au numérique dans les services d'archives. En effet, nos entretiens nous ont permis d'établir un constat : les archivistes

interrogés ont tous suivi un parcours similaire, en termes d'études et de formation, avec une formation initiale en Histoire ou à l'école des Chartes, complétée éventuellement par un diplôme (le plus souvent un Master) en archivistique. Peut-être y a-t-il également un effet générationnel, et la vision des décideurs, notamment des directeurs de services d'archives, peut elle aussi constituer un facteur déterminant pour expliquer les difficultés liées à la transition numérique. C'est en tout cas ce que semble suggérer Pascal Geneste, en mettant en avant la difficulté la transition de la collecte papier à la collecte d'archives numériques :

« On s'aperçoit quand même, 10-15 ans après, qu'on reste très largement sur cette archivistique traditionnelle, et que ceux qui étaient chargés de la collecte de certaines administrations publiques continuent à l'être sans pour autant être impliqués dans de la collecte de documents électroniques. »

Concernant la capacité des archivistes à évoluer vers une pratique courante de l'archivage électronique, Amélie Antoine admet : « le mieux pour moi, ce serait d'avoir un projet où je pourrais aller jusqu'au bout. [...] par contre, il faut que je puisse le faire quotidiennement, et pas juste une fois tous les six mois ». Dès lors, il semble difficile pour les archivistes de s'immerger dans de tels projets s'ils ne sont pas au cœur de leurs fonctions et il ne peut que difficilement s'agir d'éléments en périphérie de leur activité quotidienne. Par conséquent, des freins à cette potentielle acculturation subsistent très clairement. En effet, aujourd'hui encore, la gestion des métadonnées n'est pas considérée comme le cœur de métier des archivistes, qui veulent évidemment avant tout promouvoir la préservation d'un patrimoine et en garantir l'accès dans des conditions idéales, et ne constitue pas une « accroche » ou un point d'intérêt pour une grande partie des professionnels. Il semble alors indispensable de prouver l'interdépendance de ces techniques avec les enjeux que les archivistes ont à cœur. Ainsi, d'après Agnès Vatican, « il ne faut pas penser qu'on va embarquer toute une profession avec ce genre d'outils. » Les propos tenus par Amélie Antoine permettent également de souligner que les principes qui sous-tendent le métier des archivistes ne sont évidemment pas remis en cause par ces évolutions technologiques : « Mon métier c'est quand même de permettre la sauvegarde de la mémoire et l'accès aux sources. À terme, moi ce que je veux c'est préserver et sauver des documents de l'oubli, de permettre l'accès, de permettre la conservation de la mémoire. »

Pour autant, l'archivage électronique, nous l'avons vu, est bel et bien considéré comme une préoccupation centrale et ce n'est pas la nécessité de prendre en compte ces mutations qui est remise en question, mais plutôt les ambitions parfois démesurées et la rapidité de ces évolutions qui semblent être sources de scepticisme pour les archivistes, qui doivent désormais travailler main dans la main avec les informaticiens sur ce type de projets.

II. Le nouveau paradigme de la gouvernance de l'information : vers davantage de transversalité

II.1. L'essor de la collaboration interprofessionnelle, source de difficulté ou facteur d'enrichissement ?

Les choix qui ont été faits en termes de gouvernance, notamment en ce qui concerne la répartition des acteurs et l'organisation des différents comités qui permettent la mise au point des outils nécessaires au fonctionnement du SAEM sont cruciaux pour comprendre en quoi ce projet s'inscrit dans une logique de gouvernance de l'information. Ainsi, dans chacun de ces comités, qu'il s'agisse des instances décisionnelles ou de l'équipe opérationnelle qui est en première ligne, avec les prestataires, les différentes typologies de professionnels et les structures qu'ils représentent occupent une place leur permettant de faire valoir leurs compétences propres et les enjeux qui dépendent de leur périmètre d'action.

De fait, le comité de pilotage (COPIL), composé du directeur général des services, des directeurs des services d'archives et des systèmes d'information ainsi que des chefs de projet archives et informatique se réunit lorsque des étapes clés sont franchies et a pour mission de rendre les arbitrages nécessaires à la conduite du projet et la validation de ces étapes. Le comité technique (COTECH), quant à lui, est censé se réunir au moment des étapes intermédiaires et pour faire le point sur les avancées du projet, avec pour mission la préparation des comités de pilotage et la validation de ces étapes. Mais c'est surtout l'organisation du comité projet (COPROJ) qui est particulièrement parlante pour étudier cette transversalité et l'importance qu'elle a aux yeux de l'équipe du SAEM. Le COPROJ, dont les missions sont de réaliser et suivre l'ensemble des études et actions nécessaires à la conduite du projet et de rendre compte aux instances

décisionnelles précédemment évoquées, se réunit deux fois par semaine et s'articule autour d'un binôme archiviste / informaticien pour chacune des collectivités. Nos différents entretiens mettent en valeur l'importance de ce mode d'organisation. Ainsi, Agnès Vatican considère que la diversité des cultures professionnelles et le degré d'interaction entre informaticiens et archivistes « a vraiment été une réussite de ce projet, [avec] un groupe humain qui s'est engagé dans une piste un peu nouvelle, avec pour beaucoup une montée en compétences, voir une acculturation sur le sujet. » L'expression « montée en compétences » et le terme d'« acculturation » démontrent bien les enrichissements réciproques qui ont pu découler du mode de gouvernance mixte qui caractérise le comité projet. Lorsque nous lui demandons si ces différentes visions ont pu compliquer les échanges, elle ajoute : « J'ai plutôt ressenti qu'elles les enrichissaient. » C'est un avis qui est partagé par l'ensemble des personnes avec lesquelles nous nous sommes entretenues, puisque Sylvain Thénault rappelle lui aussi que « c'était aussi l'intérêt de cette équipe projet que d'avoir ces différents profils. »

Néanmoins, l'intensification des relations interprofessionnelles représente une source de complexité pour les différents acteurs impliqués, plaçant les archivistes dans un système qui dépasse le cadre des compétences qui leur sont habituellement assignées. Amélie Antoine présente cette problématique de la manière suivante :

« Il y a beaucoup d'acteurs, par contre, sur ce type de projets, ça c'est vrai. Avant, on pouvait classer un fonds tout seul, à deux, trois personnes, on pouvait se l'approprier, même un tableau de gestion. Là, c'est des projets qui nécessitent beaucoup de monde autour de la table. »

Cette réponse rejoint l'idée d'un nouveau paradigme en termes de gouvernance, puisqu'en raison de la multiplication des acteurs qui sont à même d'intervenir sur ces questions, il convient de mettre en place des processus leur permettant d'interagir de manière efficace. Il n'est plus possible, dans le cadre de projets qui s'inscrivent dans un tel contexte, de travailler de manière isolée. En raison de ces nombreux facteurs qui doivent désormais être pris en compte, Pascal Geneste considère, en raison de ces nombreux facteurs qui rentrent en ligne de compte, que « ce ne sont pas des systèmes très simples à modéliser », traduisant ainsi une difficulté à faire émerger des modèles communs et à mettre en œuvre des solutions transposables aux différents contextes. De manière plus concrète, c'est souvent au niveau du vocabulaire employé que cette

complexité se matérialise, comme le souligne Amélie Antoine :

« Après, la difficulté que je rencontre, c'est le vocabulaire [...] [les informaticiens] vont utiliser des termes souvent anglophones. [...] C'est plutôt en termes de terminologie que c'est pas toujours évident pour moi pour le moment, [...] le temps que je comprenne, je n'ai pas le temps d'en placer une. »

Cette réponse, donnée de manière très spontanée, reflète la nécessité d'aboutir à un vocabulaire commun, élément central de la gouvernance du projet SAEM. Le vocabulaire technique doit pouvoir être appréhendé par les archivistes, tout comme les termes spécifiques au versement et à la conservation de documents d'archives doivent être clairs pour les informaticiens. Dans cette perspective, c'est un véritable dialogue qui doit être mis en place entre ces acteurs qui sont amenés de façon de plus en plus étroite sur des projets communs. Damien Duzert, qui est amené à travailler avec différents professionnels de la gestion de l'information, confirme ce point de vue, tout en rappelant qu'il ne s'agit pas pour autant de se substituer aux acteurs en question : « Le vocabulaire est important. Il faut le comprendre effectivement, mais c'est juste de la compréhension, j'ai pas besoin d'aller plus loin [...] [c'est] la première couche, quoi. Après, je ne suis pas bilingue EAD par exemple. »

L'objectif de la mise au point d'un vocabulaire partagé est de permettre aux différents acteurs impliqués dans un tel projet de faire bénéficier à l'ensemble de l'équipe de leurs compétences respectives, sans que cela ne soit source de confusion. En effet, Amélie Antoine reconnaît que les archivistes bénéficient réellement de la vision que les informaticiens ont de l'archivage : « On a de nouveaux acteurs, de nouveaux collaborateurs, c'est-à-dire des informaticiens, effectivement, comment récupérer le bon format, comment s'assurer que ces informations sont complètes, pérennes. »

Dès lors, et nous avons pu l'observer dans le cadre de notre objet d'étude, lorsque cette vision commune émerge et que les professionnels s'accordent sur la terminologie employée, la diversité de ces cultures professionnelles peut venir enrichir considérablement la mise en œuvre de tels projets. Delphine Jamet, par exemple rappelle que « l'organisation du COPROJ a été essentielle, [...] il y a eu une vraie volonté de travailler en commun » et dresse le bilan suivant, qui est représentatif des différents retours que nous avons eu et de ce que nous avons observé :

« Je pense qu'on a eu des informaticiens qui ont très vite compris le métier des

archives, [ils] se sont mis à essayer de bien comprendre les archives et nos problématiques, avec toujours une écoute réciproque, pour que le côté technique ne l'emporte pas sur le côté fonctionnel »

C'est cette réciprocité qui constitue ici le maître mot, puisque chaque type de professionnel doit être en mesure d'apporter sa plus-value et doit pouvoir mettre en valeur ses compétences de manière optimale. Ainsi, Amélie Antoine rappelle que « [les informaticiens] m'expliquent ce qui est faisable et ce qui ne l'est pas » tout en rappelant, toujours au sujet de ce rapport privilégié entre l'archiviste et l'informaticien, que « l'un ne va pas sans l'autre par contre. » Les interactions entre archivistes et informaticiens sont donc, d'après les témoignages recueillis, fondamentales pour la réussite d'un projet de conception d'un SAE. Il doit y avoir une volonté réciproque de développer une vision et même une culture communes, et d'adapter à la fois la terminologie employée et les procédures métier afin de faire coïncider les différentes approches. Dans le cas du projet du SAEM girondin, le fait que les informaticiens se soient rapidement appropriés la manière de travailler des archivistes et l'organisation du comité projet en binômes informaticien/archiviste, ont été des points cruciaux de la conduite du projet. Le fait d'avoir un « product owner », selon l'expression consacrée, qui ait cette double casquette d'archiviste et informaticien, est également un facteur non négligeable, ce qui ressort à travers plusieurs de nos entretiens. Agnès Vatican insiste d'ailleurs sur ce dernier point : « il y avait peut-être des gens un peu devant, comme Pascal Romain par exemple, mais malgré tout il y avait un assez bon équilibre ». Le positionnement de Pascal Romain semble donc avoir été un élément clé du dialogue entre les différents acteurs du projet, comme le signale également Delphine Jamet : « On a eu la chance d'avoir Pascal Romain, qui avait la double casquette, d'attaché de conservation et en même temps d'ingénieur. » En vertu de ces observations, il nous paraît important de nous poser la question de la place d'un intermédiaire, d'un médiateur, au centre du projet pour faire émerger cette gouvernance transversale et favoriser l'efficacité des échanges.

Le fait d'avoir une équipe pluridisciplinaire permet également de faciliter le dialogue avec les prestataires, tout en s'assurant que les différentes attentes sont bien comprises par l'ensemble de l'équipe, comme l'explique Sylvain Thénault : « C'est aussi parce qu'il y a eu une forme de convergence technique entre les membres du COPROJ

et nous. À la fois, eux convergent vers des aspects plus techniques et nous on convergeait vers des aspects plus fonctionnels. »

Il s'agit toutefois de nuancer ces observations, puisque les différentes catégories de professionnels demeurent maîtres du périmètre propre à leur spécialité et la répartition des tâches peut parfois s'effectuer de manière instinctive, sans qu'elle ne pose réellement question, alors qu'elle aurait pu être sujette à discussion. Dans le cadre du projet du SAEM girondin, la place de la DSI s'est accentuée au fur et à mesure de la montée en compétences des informaticiens sur les aspects concernant l'archivage électronique et avec le lancement des phases plus techniques du projet :

« Je pense qu'au début, peut-être que les archivistes étaient plus présents mais dès que les informaticiens montent en compétence côté archives [...] c'est vrai qu'ils peuvent nous remplacer... enfin, pas nous remplacer mais en tout cas se substituer à nous, [...] ils prennent plus de poids parce qu'on est vraiment dans la phase plus technique, tout ce qui est mise en production, stabilisation des outils »

Pourtant, d'après les professionnels que nous avons pu interroger, du moins dans le cadre du projet SAEM, l'équipe aurait tout intérêt à maintenir ce dialogue pour toutes les phases du projet. Delphine Jamet s'interroge justement sur cette évolution : « Avant on réfléchissait beaucoup plus, on faisait des ateliers, là on n'a pas fait un atelier depuis 6 mois. » Ainsi, à partir d'un moment, il y a un glissement qui s'effectue de manière quasiment « naturelle » et le rôle des archivistes s'amenuise lorsque les aspects techniques occupent une place prépondérante. C'est par ailleurs un constat qu'elle dresse, non pas seulement à l'échelle du SAEM mais en ce qui concerne l'ensemble des projets qui gravitent autour du cycle de vie des documents.¹⁹⁴

Ces éléments rejoignent l'idée selon laquelle les archivistes peuvent parfois être écartés de certains aspects sur lesquels leur expérience aurait pu constituer une valeur ajoutée considérable, au profit d'informaticiens, de « techniciens », qui ont une autre vision des questions de gestion documentaire. Ainsi, Sylvain Thénault réalise lui aussi le potentiel de ce dialogue et de ces retours permanents entre les différentes entités : « Ils nous demandaient des évolutions plus ou moins grosses et du coup ça menait à des discussions en atelier avec eux pour voir la direction qui était raisonnable de prendre,

194 « On avait aussi demandé à être pilotes pour la mise en place de la GED transverse, pour la direction, parce qu'il valait mieux qu'on sache il fallait l'utiliser pour nous-mêmes [...] et on ne l'a pas été [...] pour la GED transverse, il aurait fallu un binôme bien défini aussi, avec une organisation type COPROJ. »

aussi bien d'un point de vue technique que d'un point de vue fonctionnel. » Nous le voyons, lors des phases de conception initiales, du moins avant la préparation de la mise en production, des échanges réguliers sous la forme d'ateliers ont été nécessaires afin de faire avancer le projet dans la bonne direction et d'assurer le développement d'une vision commune. C'est notamment pour préserver un certain équilibre entre les différentes visions des problématiques abordées que ces échanges doivent être maintenus. Il s'agit en effet d'éviter que la partie technique ne l'emporte sur la vision fonctionnelle des outils à mettre en place, d'autant que ces professionnels n'ont pas nécessairement les mêmes enjeux en tête : « C'est-à-dire que l'informaticien va stocker,[...] Il parle de poids, il stocke [et] c'est tout. À la limite il va stocker tout ce qu'il peut, il va stocker, il va stocker. Mais on ne lui demande pas toujours de stocker. » Dans cet exemple, évoqué par Amélie Antoine, les informaticiens pensent davantage en termes de stockage et de sécurité des systèmes d'archivage que de pérennité et de valeur probante du document. Les aspects techniques et les considérations liées à la volumétrie peuvent alors prendre le pas sur les enjeux qui sont au cœur des métiers des archives.

II.2. Gouvernance et gestion de projet, quelle imbrication et quelles conséquences pour l'archiviste ?

Méthodologie de gestion de projet et rôle des décideurs : quel impact sur la gouvernance ?

La présentation du projet SAEM effectuée aux RASAD¹⁹⁵ en février 2017 rappelle les choix qui ont été effectués en matière de gestion de projet, à savoir notamment un recours aux méthodes agiles pour le développement, avec des « sprints » (c'est-à-dire des itérations) sur trois ou quatre semaines et une déclinaison en « tâches de développement ». Cette orientation n'est pas anodine puisqu'elle implique également une rédaction en commun des « user stories », à laquelle les archivistes impliqués dans le projet participent évidemment. Cette manière de procéder, qui a pour objectif de faciliter le dialogue avec le prestataire, implique une certaine acculturation pour les archivistes, n'étant pas monnaie courante au sein des services d'archives.

Pour répondre à cette problématique de conduite du changement, les besoins en

195 Rencontres annuelles de la section des archivistes départementaux

termes de formation à ces nouvelles formes de gestion de projet ont fait l'objet d'une attention particulière dans le cadre du développement du module référentiel, et ce dès la rédaction du cahier des clauses techniques particulières. Ainsi, y sont inscrites les demandes suivantes vis-à-vis du prestataire : « réalisation d'un séminaire de sensibilisation aux possibilités et enjeux techniques », « formation à l'utilisation de l'outil de gestion du backlog » et « construction, en lien avec le groupe projet d'un backlog. »¹⁹⁶

Il s'agit donc d'une manière de concevoir l'organisation d'un projet qui implique de développer une certaine vision, voire des compétences particulières. La transition vers le « mode projet », d'après les dires de nos interlocuteurs, avait déjà posé le même type de questionnements au sein des Archives départementales, comme le souligne Agnès Vatican : « Pour [les informaticiens], je dirais que ce fonctionnement en chef de projet était plus habituel, pour les archivistes c'était peut-être plus nouveau. » Cela laisse supposer que la différence en termes de culture professionnelle et cette émergence en termes de besoins de gouvernance ne se fait pas ressentir que sur le plan technique mais peut s'observer également au niveau de l'organisation des équipes et la manière de travailler. Sylvain Thénault, qui a pu voir l'évolution de l'équipe au fur et à mesure de l'avancée du projet, confirme ces éléments :

« Clairement, il y a des gens qui avaient différents points de vue sur les aspects techniques et qui du coup ne partent pas tous du même niveau, il y a des choses qui sont plus ou moins faciles à appréhender pour certains que pour d'autres, et ça se voyait aussi dans les réunions de gestion de projet, de manière générale, pas que sur les choses techniques, mais également fonctionnelles »

Par ailleurs, dans un processus de gouvernance de l'information, l'implication des décideurs au sein des projets liés à la gestion de l'information et des données est primordiale. Nos entretiens avec Pascal Geneste et Agnès Vatican, membres des COTECH et COPIL précédemment présentés, nous permettent alors d'interroger cette facette du sujet. Pascal Geneste rappelle alors son expérience personnelle : « J'ai été tout de suite intégré au comité de pilotage pour travailler sur la première phase d'étude » et précise par la suite que son rôle a notamment consisté à « participer au comité de pilotage et au comité technique lorsqu'ils se réunissent, et d'essayer de

196 « Ensemble de fonctionnalités détaillées comprenant la définition des user stories, de leur complexité, les éléments de priorisation et de définition du prêt à être discuté. »

promouvoir à l'extérieur le projet, notamment en essayant de lever des subventions auprès du SIAF chaque année, en participant à l'appel d'offre AD-ESSOR » ainsi qu'à « la mise en place de la convention ». Ces éléments soulignent un positionnement clé des décideurs, mais indiquent pour autant que leur rôle est souvent limité à la mise en œuvre de conventions de partenariat et aux échanges avec les structures extérieures, dans une optique de valorisation du projet. La directrice des Archives départementales de la Gironde confirme ces éléments, en rappelant leur importance dans le cadre du dialogue avec les élus :

« Mon rôle, c'est de valider la stratégie, la feuille de route, voire parfois même demander des inflexions dans le projet, voilà. Et puis faire l'articulation, avec les élus bien sûr, puisqu'on a aussi ce rôle, les directeurs et les DGA, d'articulation avec le champ politique et voir comment on répond aux attentes des politiques »

Un tel positionnement des instances décisionnelles, qui ne participeraient pas nécessairement à l'élaboration des documents stratégiques, qui sont le fondement d'une gouvernance de l'information, mais simplement à leur « validation », pourrait alors constituer un facteur limitant le développement d'une stratégie globale.

Du référent au spécialiste : la question du positionnement d'acteurs clé au sein même des services

Dans la partie précédente, nous soulignons l'inquiétude de certains de nos interlocuteurs quant à la capacité des archivistes à s'approprier rapidement les outils développés dans le cadre du SAEM girondin. Cette difficulté peut être due, comme nous l'avons déjà évoqué, à une mise en production qui tarde à se mettre en place, mais plus largement à la manière dont des aspects tels que l'archivage électronique et la gestion de métadonnées sont perçus au sein même des services d'archives. En parlant de Delphine Jamet, son adjointe, Pascal Geneste admet qu'il lui laisse prendre les devants sur ces questions : « son expérience, depuis une quinzaine d'années, de développement de systèmes d'information, on va le dire rapidement comme ça, me permet de m'appuyer sur quelqu'un d'extrêmement solide et qui comprend bien la problématique. » En effet, les archivistes qui ont développé ce type de compétences sont souvent considérés comme des spécialistes et il est alors possible de retrouver, au sein des équipes, des agents référents en matière d'archivage électronique, sur lesquels les décideurs, tout comme les autres agents, peuvent s'appuyer.

Naturellement, nous pouvons nous reposer sur les propos tenus par l'intéressée pour compléter cette affirmation :

« J'avais été chef de projet pour l'outil de mise en ligne, Pléade, pour GAEL. Donc là, pareil, j'ai été amenée à travailler de façon étroite avec la DSI, et les prestataires bien sûr, qui développent, [...] c'est plutôt tout mon côté spécialiste des métadonnées, des instruments de recherche, etc. qui m'a amenée dans le projet. »

La place d'un spécialiste des questions liées à l'archivage électronique et aux problématiques connexes, qui soit en mesure de dialoguer avec les services informatiques et les prestataires, est donc primordiale. Ce constat peut être effectué par le biais du choix des archivistes qui ont été intégrés au comité projet du SAEM, puisque Agnès Vatican admet que ces derniers sont « plutôt des archivistes qui étaient sur de la collecte d'archives contemporaines, donc qui avaient quand même un peu tous cette orientation. » D'après cette réponse, nous pourrions supposer qu'il existe des profils, parmi les archivistes, plus à même de dialoguer avec les autres typologies de professionnels, favorisant alors l'émergence, plus ou moins spontanée, de spécialistes au sein des services d'archives. De la même manière, Damien Dudezert, en tant que responsable applicatif pour la gestion documentaire du Département de la Gironde, est lui aussi en mesure d'observer l'importance de ces référents, soulignant ainsi que cette acculturation est loin d'être généralisée : « Alors, il y en a qui maîtrisent, on sent que quand on parle avec Delphine, ou même Pascal, il y a une maîtrise sur le sujet [...] il y en a qui perçoivent ce parcours de données [et] avec certains profils, on discute de manière un peu plus approfondie sur le sujet, [...] j'ai trois ou quatre personnes référentes aux Archives. » En effet, les archivistes qui se saisissent de ces questions ne sont que peu nombreux, bien que l'ensemble de la profession y soit de plus en plus confrontée dans la pratique. Nous l'avons vu, aux Archives départementales de la Gironde, il n'y a pas de réel spécialiste de l'archivage électronique, qui s'occuperait de ces aspects à temps plein, mais plutôt des référents qui ont développé une appétence pour ces questions au fur et à mesure des évolutions technologiques. Pour autant, c'est bien l'ensemble des archivistes qui devra être en mesure de maîtriser ces aspects-là, ce qui pose là encore la question de la mise en correspondance avec leur cœur de métier.¹⁹⁷

197 D'après P. Geneste : « Mettre des agents sur la conception d'un référentiel, [...] sur la normalisation et la structuration des données etc, c'est important, mais il faut leur expliquer pourquoi et dans quel contexte on le fait, ce que ça va donner, la mise en perspective de leur travail quotidien. »

III. Web sémantique et données archivistiques, un enchevêtrement souhaité et souhaitable ?

III.1. Harmonisation des données, automatisation du traitement des documents : le web sémantique au service des archivistes

Le module référentiel du SAEM : un outil de gouvernance au service de ses utilisateurs

L'analyse des spécifications fonctionnelles relatives à la pérennisation des métadonnées du SAEM, élaborées dès la fin de l'année 2013 avec le concours du prestataire AKKA, permet de cibler les principales attentes ainsi que les objectifs du référentiel. Ainsi, l'un des tous premiers termes permettant de caractériser l'intérêt de ce module est l'interopérabilité, l'objectif étant au final de « pouvoir conserver les métadonnées nécessaires à associer aux fonds, dossiers et documents archivés (papiers et électroniques) en permettant l'automatisation de leur récupération, de leur génération et de leur organisation. » En vertu de ce besoin d'interopérabilité, la question de la dépendance aux outils et de la compatibilité des logiciels et des systèmes d'information est posée par l'équipe du COPROJ de la manière suivante :

« Nous sommes actuellement confrontés à une couverture partielle des besoins en fonction de la problématique abordée (conservation technique, description archivistique, communication sur Internet, etc.) entraînant des risques d'incompatibilités ou de faible interopérabilité entre les systèmes d'information. »

C'est pour assurer cette compatibilité et limiter l'adhérence aux différents outils qui peuvent être déployés dans le cadre d'un SAE que le recours à un environnement logiciel reposant sur certaines briques du web sémantique a été envisagé pour mettre en place ce référentiel. Nous l'évoquons lors nous présentions le contexte de notre expérimentation, le logiciel CubicWeb développé par Logilab fait reposer son interface sur un modèle de données. Ce modèle de données est le fruit d'un travail en commun entre le prestataire et l'équipe du COPROJ et a pu évoluer en fonction des mutations fonctionnelles et techniques de l'archivage électronique. C'est donc l'enjeu d'une exploitabilité sur le long terme, à travers une volonté de garantir la réversibilité des données, qui sous-tend l'élaboration du référentiel de données du SAEM girondin, comme le rappelle le cahier des clauses techniques particulières du projet :

« Par définition l'ensemble des métadonnées est mouvant et changeant au cours du cycle de vie (génération d'informations techniques dès la création, ajouts de métadonnées métiers, enrichissement de métadonnées liées à l'archivage ou aux opérations techniques de migrations de formats par exemple). Pour autant cet ensemble doit rester cohérent et exploitable pour les différents types d'utilisateurs. »

Ainsi, pour Sylvain Thénault, ce n'est peut-être que peu surprenant au vu de ses fonctions et son positionnement dans le projet, l'intérêt du recours aux différentes couches du web sémantique est avant tout lié à des raisons technologiques, dans la mesure où il permet de limiter l'adhérence à des outils ou des formats spécifiques, en passant par l'intermédiaire de modèles de données et de vocabulaires interopérables :

« L'avantage se résume, pour moi, en un mot, on va dire principal, c'est l'interopérabilité. Donc être capable de faire des systèmes qui marchent ensemble plutôt que la démarche opposée, à savoir faire un gros système qui retient tout. »

En outre, les spécificités liées au traitement d'archives, en raison de l'importance des notions de provenance et de contextualisation de la production documentaire, engendrent des impératifs auxquels l'outil, à travers les métadonnées qui y sont stockées et exposées, doit répondre. Le référentiel doit effectivement être en mesure, comme le signale le cahier des clauses particulières, d'assurer la préservation du contexte de production¹⁹⁸, de gérer l'ensemble des accès et des droits liés au régime de communicabilité des archives¹⁹⁹, de garantir l'intégrité de l'information²⁰⁰, mais surtout de mettre en œuvre la mise en relation de toutes ces données issues de l'archivage.²⁰¹

Toujours d'après le CCTP du module référentiel, et pour assurer son interopérabilité avec l'ensemble des normes qui gravitent autour de la gestion des documents numériques, le référentiel de données « doit être pensé comme un système d'organisation de connaissances, basé sur une ontologie et permettre l'automatisation de la récupération et l'organisation des métadonnées des productions électroniques. »

198 Par l'intermédiaire d'un lien vers la description de l'auteur du document mais également celle du contexte administratif ou privé au sein duquel il a été produit.

199 En effet, d'après le CCTP : « les habilitations liées aux documents évoluent au cours du cycle de vie de l'information et cette information doit être préservée pour garantir les aspects réglementaires liés aux conditions de leur diffusion. »

200 « Même sous a forme dématérialisée, l'information documentaire dépend des supports et des formats de fichiers dans lesquels elle est stockée et structurée. Il s'agit donc d'intégrer ces contraintes aux métadonnées afin de garantir leur intégrité. »

201 En vertu de la manière dont s'effectue le classement en archivistique, chaque document doit pouvoir être rattaché à un ensemble significatif pour qu'il soit possible d'en saisir le contexte et d'en appréhender les différentes dimensions. Au-delà du contexte de production, il est également nécessaire de documenter son lieu de production et de rattacher son contenu à des termes d'indexation qui permettent de le classer et de conserver sa dimension temporelle.

Il doit également fournir, pour réduire l'adhérence aux couches applicatives, un cadre de référence pour la structuration de l'information de pérennisation utile aux gestionnaires du SAE.²⁰² Pour garantir son interopérabilité et assurer le suivi de l'outil sur le long terme, mais également dans l'optique d'une ouverture vers des dispositifs similaires et à l'exposition des données sur le web, le module référentiel a également été conçu pour être capable d'identifier les métadonnées de provenance propres à une collectivité, de dédoubler les données de référence importées depuis les sources de données externes mais aussi d'interroger des référentiels externes, disponibles sur le web, afin de permettre d'éventuels enrichissements ou alignements.

En février 2017, lors de la présentation du dispositif aux RASAD, les objectifs sont restés les mêmes, puisqu'il est toujours question de « créer, importer, centraliser, gérer et exposer des vocabulaires contrôlés publics ou internes des notices d'autorité des acteurs de l'archivage, des autorités et unités administratives ainsi que des profils d'archivage conformes au SEDA » dans l'optique d'une « simplification des processus d'archivage par réutilisation des données sans re-saisie (synchronisation). » Nous le voyons, l'harmonisation et la réutilisation des données sont les principaux moteurs du module référentiel et justifient, de ce point de vue, le recours aux technologies du web sémantique, dans la mesure où certains de leurs mécanismes vont permettre d'identifier et de désambigüiser les données. Les user stories rédigées par le comité projet à l'occasion d'un atelier mené avec la société Anaphore au sujet de l'interconnexion avec le SIA utilisé par les Archives départementales de la Gironde vont également dans ce sens²⁰³ et sont symboliques des attentes des archivistes, qui exigent avant tout un allègement des tâches les plus répétitives, qui sont également porteuses de risques liés à l'erreur humaine, lors de l'indexation par exemple.

Pour mettre en corrélation la manière dont sont pensées ces différentes couches qui sont l'essence même du module référentiel et la manière dont leurs enjeux sont perçus par les archivistes, nous avons interrogé ces derniers au sujet de la place qu'occupe cette brique dans l'ensemble du dispositif d'archivage électronique girondin. Ainsi, Pascal Geneste considère que « c'est une place essentielle, [...] le référentiel, pour moi, c'est le module phare, de tout système, qu'il soit traditionnel ou beaucoup plus

202 Cf. liste des fonctionnalités attendues du référentiel en annexe 8.

203 Cf. users stories 1 à 3, en annexe 16.

innovant. » Delphine Jamet, quant à elle, le définit de la manière suivante :

« Un seul outil qui permet d'éviter de toujours répéter les choses, d'avoir tout en un seul endroit, que quand on cherche un mot, on sait que c'est là qu'on va le trouver et pas ailleurs [car] on gère beaucoup de noms, entre les noms de producteurs, les noms géographiques, les noms de personnes, les noms de notaires, les mots matière etc, on en gère énormément. »

D'après elle, « [le référentiel] peut vraiment être pivot pour qu'on ait les mêmes termes d'indexation, en tout cas qu'il y ait des alignements, et du coup que nous on aura beaucoup moins de travail à faire à terme sur les archives électroniques qu'on va conserver. » Nous le voyons d'après les témoignages recueillis, qui confortent les intentions initiales que nous avons précédemment soulevées, le module référentiel occupe une place essentielle dans l'organisation du projet du SAEM girondin, puisqu'il permet avant tout d'harmoniser les données qui y sont stockées. Par extension, il permet de faciliter le travail des archivistes, qui auront des ressources normalisées et centralisées pour mener à bien le traitement des versements et l'indexation des documents. Il s'agit également d'un outil leur permettant de créer, de gérer et d'exporter l'ensemble des notices producteur en EAC, en restant sûrs et certains que les données renseignées sont exploitables. L'expression clé est ici « moins de travail à faire », il s'agit véritablement d'épargner aux archivistes des tâches fastidieuses et répétitives, tout en garantissant la qualité des informations. En outre, c'est vraiment cet impact sur les tâches quotidiennes qui est mis en avant par Delphine Jamet, qui considère que les archivistes sont et doivent être les utilisateurs principaux d'un tel outil.²⁰⁴

Le recours à des identifiants uniques et pérennes

Comme le mettait en lumière le CCTP du référentiel, l'attribution d'identifiants aux différents types de données qui sont stockées au sein du référentiel constitue l'une des composantes principales de cet outil transversal. Ils constituent le socle de l'architecture de CubicWeb et sont le reflet direct de la place des URI dans le fonctionnement même du web sémantique, tel qu'il a été pensé dès son origine. À ce niveau, les équipes du projet SAEM ont rapidement pris la décision de s'orienter vers

²⁰⁴ « Pour moi, les premiers qui vont l'utiliser, ça va être les archivistes, ça ne va pas être les services. Je pense que c'est une divergence qu'on peut avoir niveau SI, enfin entre SI et archivistes, parce qu'effectivement ce ne sont pas les services qui vont tout de suite l'utiliser, ça moi j'en suis persuadée. »

un modèle similaire à ce que propose le portail Data BnF, avec l'intégration d'autorités nommantes, qui permettent d'attribuer des identifiants ARK à différents types de ressources et donc d'assurer un certain contrôle et une continuité dans leur syntaxe.²⁰⁵ Cette volonté s'inscrit dans la politique de données de références développée à l'échelle du SIAF et du ministère de la Culture, qui cite notamment dans son vade-mecum pour les producteurs de données²⁰⁶ le recours aux identifiants uniques comme pilier d'une gouvernance de l'information.

Ainsi, dans le module référentiel, des identifiants de ce type sont attribués à l'ensemble des vocabulaires, concepts et notices d'autorité mais également aux profils SEDA, aux différentes autorités administratives et aux agents de l'archivage. In fine, c'est l'ensemble des types de données qui sont vouées à être récupérées par la GED-SAS et As@lae qui sont concernés, en vue de l'optimisation de leur réutilisation. Par ailleurs, pour faire le lien avec les problématiques propres aux archivistes, le SIAF, lors d'échanges avec l'équipe du SAEM girondin en 2014, lorsque le module référentiel était encore en phase conceptuelle, rappelait alors les difficultés en termes de cotation des documents électroniques et précisait en quoi le recours à ces identifiants pourrait constituer une solution :

« Les circulaires existantes sur les cadres de classement et modalités de cotation ne sont pas adaptées aux besoins en matière de cotation pour l'archivage électronique. Il faut privilégier l'utilisation d'identifiants pérennes de type Ark dans les systèmes d'archivage électronique. Sur ce point, une information à l'ensemble du réseau est envisagée par le SIAF. »

De fait, le recours à des identifiants pérennes dépasse le simple cadre du référentiel de données et semble pouvoir s'inscrire directement dans les pratiques des archivistes, en se substituant alors aux modalités actuelles en termes d'identification et de localisation des ressources électroniques.

Notre analyse des différents tickets issus du backlog mis en place dans le cadre de la

205 D'après le manuel en ligne du SAEM : « un ARK est un identifiant pérenne d'une ressource ou d'un objet disponible sur le web, quelle que soit sa nature et indépendamment de modifications du site sur lequel elle/il apparaît. » Le système ARK, conçu par la California Digital Library, étant donné qu'il repose entièrement sur le mécanisme des URI, permet d'identifier une ressource sur le long terme, de manière pérenne. Une ressource en ligne peut alors être identifiée, indépendamment des mutations que subissent les dispositifs d'hébergement et les protocoles d'accès.

206 Vade-mecum pour les producteurs de données – Ministère de la Culture.[en ligne]. [Consulté le 12 août 2017]. Disponible à l'adresse : <http://www.culturecommunication.gouv.fr/Divers/Harmonisation-des-donnees-culturelles/Actualites/Vade-mecum-pour-les-producteurs-de-donnees>.

conception du référentiel de données nous a permis d'observer qu'une part considérable des user stories rédigées par le comité projet concernent la mise en œuvre de ces identifiants et qu'elles sont par ailleurs parmi les premières à avoir été rédigées.²⁰⁷ Les rédacteurs des tickets relatifs à ces aspects ont par ailleurs clairement exprimé, à travers plusieurs critères d'acceptation, la nécessité d'attribuer de manière automatisée un identifiant pérenne aux acteurs, aux vocabulaires et aux concepts qui s'y trouvent, au moment de leur création ou de leur import. Les critères d'acceptation ciblent les besoins d'unicité des identifiants dans la base de données du référentiel, pour les différentes unités d'archives, et de désolidariser leur traitement de leurs intitulés. Ces identifiants doivent avant tout permettre de faciliter le travail des archivistes, et de façon plus globale de tous les utilisateurs du système, en permettant d'automatiser les processus d'import et de création de données de référence.

En outre, Delphine Jamet, qui a elle-même rédigé certains de ces tickets, défend l'utilité de ces identifiants en rappelant que leur absence au sein des portails et sites nationaux à disposition des archivistes constitue une difficulté pour l'exploitation de certaines ressources. Elle cite pour cela l'exemple du site des Archives nationales : « Il n'y a pas d'identifiant ARK, donc quand ils changent de système, tout mon système plante, il faut revoir tous mes liens. » Elle ajoute ensuite que lorsque ces dispositifs sont en place, à partir de ces liens : « ça nous permet de faire moins de travail, en fait, de reprise. » C'est la problématique de l'ouverture des données et leur intégration à des dispositifs nationaux et supranationaux qui constitue le principal enjeu pour Amélie Antoine : « quand tu normalises un descripteur géographique, tu fais en sorte d'identifier... de te dire qu'aujourd'hui, c'est une plateforme internationale, qu'il faut accéder à l'information, qu'il faut pouvoir rebondir. »

Ainsi, le recours à des identifiants uniques et pérennes, évoqué ici par le biais de l'intégration d'identifiants ARK au module référentiel du SAEM girondin, semble également être perçu comme une manière de rationaliser le processus d'archivage d'un document numérique, puisqu'il permet de garantir que les liens, notamment entre les vocabulaires utilisés pour l'indexation, soient établis. Pour Sylvain Thénault, c'est par ailleurs l'un des aspects les plus fondamentaux du fonctionnement du référentiel.

207 Cf. user stories 4 et 5 en annexe 16.

D'après lui, ce sont ces premières couches du « layer cake » du web sémantique qui présentent le plus de potentiel, étant les plus essentielles pour permettre l'interopérabilité des données à une échelle dépassant le système pour lequel elles sont conçues : « On essaie par exemple d'avoir des identifiants, par exemple, qui sont globaux, et qui dépassent le cadre d'une seule application. »

Vocabulaires contrôlés et ontologies, désambiguïser les données pour favoriser l'interopérabilité

Les spécifications fonctionnelles relatives à la pérennisation des données du SAEM girondin, établies elles aussi lors de la première phase du projet, rappellent par ailleurs l'importance du recours à des vocabulaires contrôlés, autre couche essentielle du module référentiel, bénéficiant des apports du modèle RDF :

« Au-delà de ces normes de description, la question centrale est celle des vocabulaires et de leur contrôle. Le groupe de travail Library Linked Data a identifié et recensé de nombreux candidats fournissant des listes d'autorités ou vocabulaires contrôlés qui permettent d'exprimer en RDF les éléments nécessaires à la structuration de l'information au sein du processus d'archivage. »

Dans cette perspective, et tout comme pour la mise en œuvre des identifiants ARK, il n'est pas surprenant de voir que les user stories rédigées par le COPROJ mettent en valeur, de par leur fréquence et la proportion qu'elles représentent, la place qu'occupe la mise en place de vocabulaires contrôlés au sein du module référentiel.²⁰⁸ Les critères d'acceptation, qui précisent la mise en œuvre technique de ces vocabulaires, insistent sur la possibilité d'importer, via des formulaires, des thésaurus au format SKOS ou des listes d'autorité en CSV. Dès le début de l'année 2015, lorsque la phase de conception du module référentiel a débuté l'import de vocabulaires SKOS a occupé une place de premier plan, notamment avec l'intégration de thésaurus externes permettant d'interconnecter le SAEM avec d'autres référentiels, avec ici l'exemple du thésaurus du ministère de la Culture relatif au traitement des archives locales :

« On veut pouvoir importer des thésaurus au format SKOS XML. Critère d'acceptation : être capable d'importer le thésaurus SIAF du ministère de la culture. La provenance des concepts doit pouvoir être tracée (par ex. grâce à la notion de source de données de cubicweb). »

La mise en place de vocabulaires en SKOS répond également à une volonté d'interconnexion des données, non seulement au sein du module référentiel mais

²⁰⁸ Cf user stories 6 à 8 en annexe 16.

également entre les différentes applications du SAEM. Comme le démontrent plusieurs exemples de tickets rédigés à cette occasion, il s'agit d'enrichir les différents contenus et de permettre, notamment à la GED-SAS, de récupérer les concepts issus des vocabulaires afin de caractériser les versements.²⁰⁹

L'intégration de ces vocabulaires de manière normalisée et interopérable, à travers le recours au SKOS, répond également à un besoin d'interconnexion avec le SIA des Archives départementales. Ainsi, lors d'un atelier dédié à cet interfaçage ayant eu lieu en décembre 2016, le comité projet et l'équipe d'Anaphore l'envisageaient de la manière suivante :

- « Il faudra donc faire des exports des index, [...] les récupérer en csv et les importer dans le référentiel pour pouvoir les exporter en skos depuis le référentiel vers Arkhéia où ils sont stockés dans une table pour être utilisés dans le module aide au classement. »
- « Dans le sprint 2, développement d'une interface pour consommer les vocabulaires dans aide au classement ou n'importe quel autre module d'Arhéia, en sélectionnant le type de vocabulaire (ex : geoname ou autre), puis, le vocabulaire choisi (un vocabulaire spécifique), puis un concept. »

L'objectif du recours aux vocabulaires contrôlés en SKOS est alors multiple, puisqu'il s'agit non seulement de standardiser les termes utilisés pour l'indexation des versements effectués via le SAEM, mais aussi d'automatiser l'indexation à partir du logiciel métier des archives, tout en permettant de tisser des liens avec des vocabulaires externes.

Un autre élément essentiel de l'intégration des technologies du web sémantique au dispositif girondin est le recours à des ontologies permettant l'indépendance par rapport aux outils utilisés et leur intégration au modèle de données qui sous-tend l'architecture du module référentiel. Nous mentionnions dans la partie théorique de ce travail l'intérêt, pour les services d'archives, du recours à une ontologie telle que PROV-O pour décrire la provenance et contextualiser les ressources. C'est, dans le cadre du SAEM girondin, la solution qui a été retenue pour assurer la prise en compte de la provenance des données et suivre les évolutions du module référentiel, à travers des métadonnées de gestion, notamment par le biais des entités prov : activity et prov : maintenanceEvent de l'ontologie.²¹⁰ Ainsi, l'intégration de PROV-O est prévue dans

209 Cf. user stories 9 à 11 en annexe 16.

210 Cf. user stories 11 à 13 en annexe 16.

l'optique d'assurer la traçabilité des données qui sont intégrées au référentiel, en permettant la gestion de leur cycle de vie. En cela, elle contribue clairement à répondre aux impératifs liés à l'administration électronique et aux contraintes légales auxquelles sont soumis les documents à valeur probante. De fait, elle contribue à répondre aux besoins en termes de preuve (« proof ») et de confiance (« trust ») situés au sommet du Semantic Web Stack.

Les notices d'autorité et les profils SEDA pour encourager la réutilisation et faciliter les échanges

Si les vocabulaires contrôlés permettent de construire des données de référence et de structurer celles-ci au sein d'ensembles qui font sens pour l'indexation des documents, les principaux outils de la réutilisation des données au sein du système girondin se présentent sous deux formes : les notices d'autorité²¹¹ et les profils SEDA. Pour Delphine Jamet, qui a contribué au développement de ces outils normalisés auprès des agents du bureau collecte et traitement des AD de la Gironde, l'objectif de l'intégration des notices d'autorité au sein du référentiel de données est clair : « On avait besoin d'un outil qui nous permette de faire des notices d'autorité propres, donc là maintenant on l'a. Le but c'était ça, de ne pas répéter les informations partout. »

Dans l'optique d'une automatisation des versements liés à un flux d'archives électroniques, l'analyse des caractéristiques de ce flux permet de définir un plan de classement type et de préciser le contenu attendu avec une normalisation des intitulés ainsi que la définition de règles de gestion qui s'appliqueront à ces archives, tels que la durée d'utilité administrative, le sort final ou la communicabilité. C'est au sein du profil SEDA, que nous avons évoqué dans notre première partie, que ces règles sont inscrites.²¹² Ces profils peuvent notamment être dupliqués afin d'être adaptés à d'autres collectivités ou services, si ces derniers visent des typologies documentaires similaires,

²¹¹ Dans le cadre du SAEM girondin, les notices d'autorités sont des fiches de description des producteurs d'archives rédigées selon la norme ISAAR-CPF et intégrées dans le référentiel à l'aide du schéma de métadonnées EAC-CPF. Elles ont pour but de décrire les autorités administratives, notamment les collectivités, en précisant leurs dates extrêmes d'existence, les éventuels changements de noms tout en indiquant leur organisation et en assurant le suivi de leurs évolutions. Elles peuvent également être employées pour décrire les unités administratives, notamment les services producteurs ainsi que leurs missions et leurs relations hiérarchiques ou d'association avec d'autres services. Les notices d'autorité permettent aussi de décrire les agents de ces autorités administratives, de préciser leur rôle et leurs fonctions, mais aussi de décrire toute personne physique ou morale produisant des archives privées se voyant confiées à un service d'archives.

faisant l'objet d'un traitement identique. Le principal objectif est donc là aussi d'harmoniser les plans de classement qui permettent de traiter les versements d'archives numériques, et par extension de réduire le temps passé à leur conception. Ces différents outils, qu'il s'agisse des vocabulaires contrôlés, des notices d'autorité ou encore des profils SEDA constituent alors, dans le cadre du système girondin, des outils visant à faciliter les tâches qui incombent aux archivistes. En tant que tels, ils trouvent tout leur sens auprès des professionnels concernés, du moins ceux qui ont d'ores et déjà pu expérimenter avec la version de l'outil en « pré-production. » Cependant, en raison de l'avancement du projet et de l'absence d'utilisateurs représentatifs, l'impact de ces technologies sur les pratiques quotidiennes des archivistes, notamment pour ceux qui ne sont pas spécialistes de l'archivage électronique, est encore flou.

III.2. Les technologies du web sémantique, quel impact sur les pratiques des archivistes ?

Des mutations technologiques caractérisées par un manque de visibilité

Si le potentiel de ces technologies en matière d'uniformisation des pratiques, à travers l'exemple du module référentiel propre au SAEM girondin, n'est en théorie plus à prouver, un manque de visibilité semble clairement régner, comme nous le supposons, sur la manière dont ces évolutions vont prendre forme dans les outils utilisés par les archivistes. Nos entretiens nous ont effectivement permis de cibler l'un des obstacles principaux à cette compréhension des enjeux du web sémantique, à savoir le vocabulaire complexe et varié qui caractérise cette conception du web. Lorsque nous tentons d'estimer le degré de familiarité de Delphine Jamet, qui est pourtant la spécialiste de l'archivage électronique du bureau auquel elle est rattachée, elle admet : « J'arrive à comprendre, en gros, mais je peux pas expliquer, c'est compliqué pour moi encore. » Ces éléments reflètent clairement une difficulté à faire la part des choses parmi les différents concepts qui gravitent autour de la transition « 3.0 » évoquée par la feuille de route du MCC²¹³, et le vocabulaire employé n'en facilite pas l'appropriation

212 Un profil SEDA peut être associé à une ou plusieurs unités administratives de type service versant qui vont pouvoir les réutiliser, dans la GED-SAS, pour effectuer un versement. C'est l'ensemble des éléments intégrés au profil et les métadonnées rattachées aux documents transférés qui vont permettre la création d'un bordereau de versement.

213 Toujours d'après Delphine Jamet : « Tout le monde parle de web sémantique et de web des données pour la même chose. »

par les archivistes. Les propos tenus par Amélie Antoine rejoignent également ce point de vue : « Parce que c'est vrai que le web sémantique, on parle tout le temps de web sémantique, c'est pas toujours évident pour moi de comprendre tout ce que ça recouvre derrière [...] je comprends ce qu'est le web sémantique mais en même temps je trouve qu'on met le web sémantique à toutes les sauces. » Pour Agnès Vatican, c'est la mise en correspondance avec les outils de recherche qui pose problème et qui contribue à un sentiment d'incertitude.²¹⁴ Par ailleurs, lorsque nous interrogeons Delphine Jamet au sujet de l'existence de projets articulés autour du web des données dans le secteur culturel, celle-ci nous répond : « Il n'y en a pas beaucoup, je crois, dans le domaine culturel. » Cette réponse, qui est représentative de l'absence de connaissance de nos interlocuteurs par rapport à de tels dispositifs, en dehors du portail FranceArchives, laisse supposer que le faible degré de participation des services d'archives à des projets qui reposent sur cette vision du web semble les couper des initiatives qui peuvent être mises en place dans les bibliothèques ou dans les musées, par exemple.

En outre, au cours de notre entretien avec Damien Duzert, celui-ci nous a rappelé que, dans le cadre de la rédaction du cahier des charges du nouveau site des Archives départementales de la Gironde, « le web sémantique, c'est ce qu'on a appelé un lot complémentaire. » En effet, malgré une réelle volonté de mettre en place une recherche fédérée extensive et de permettre une ouverture du nouveau site web des Archives départementales de la Gironde sur l'extérieur, la place qu'occupent les technologies du web sémantique dans le cahier des charges n'est que très restreinte. Ainsi, toujours dans le contexte de la création du nouveau site des Archives départementales, le manque de maîtrise par rapport à ces questions et d'une réelle visibilité du potentiel du web des données pose problème dès la mise en place des marchés et lors du processus de rédaction des cahiers des charges, puisqu'il est alors difficile d'émettre des demandes précises. La question de l'impact en termes de développement et de coûts que cela peut impliquer n'est également pas claire, comme le souligne notre interlocuteur.²¹⁵

Comme nous l'avions évoqué, pour Sylvain Thénault, l'intérêt des couches qui se

214 « A un moment, la recherche à facette semblait être l'alpha et l'oméga, aujourd'hui le web sémantique... Je ne sais pas, [...] pour moi c'est encore trop flou »

trouvent à la base du Semantic Web Stack ne fait pas de doute : « La base qui est : « on va parler avec des vocabulaires communs sur la base d'un protocole qui est aujourd'hui HTTP, en nommant les choses avec des URL », déjà rien qu'en respectant ces principes-là, qui sont finalement assez simples, on va assez loin. » D'après lui, elles seraient même très largement suffisantes pour obtenir des résultats probants en matière d'interopérabilité et d'échange des données. Pour autant, il émet certaines réserves quant à l'applicabilité des briques supérieures de ce modèle.²¹⁶ Dès lors, les doutes quant à la mise en application de ces technologies ne concernent pas seulement les archivistes, le prestataire admettant ici être moins convaincu par les briques plus complexes de ce modèle propre au web des données, dans la mesure où ce sont des éléments qui n'ont pas encore réellement abouti et qu'il ne les considère pas comme « matures ». En outre, il souligne le fait que ce n'est pas l'aspect technique qui est primordial dans le cadre de l'application de ces technologies mais plutôt la conception des modèles qui doivent permettre leur mise en œuvre, qui est l'objet même, comme nous l'avons vu, d'une gouvernance de l'information.

Pour les archivistes eux-mêmes, les retombées concrètes du potentiel du web des données sont également difficiles à entrevoir, puisque ces aspects sont encore difficilement intégrés aux projets liés aux interfaces de consultation. Pour Pascal Geneste, l'intérêt de l'intégration des données issues du traitement des archives au sein de portails fédérés sur le web relève pour lui davantage d'un gain de visibilité pour les services qui y participent :

« L'intégration de certains de nos projets à des portails nationaux ou internationaux, essentiellement européens [...] est aussi une forme de publicité du service qui participe à ce genre d'actions. Les enjeux pour une structure comme la nôtre, c'est de mieux se faire connaître d'abord, et de mieux faire connaître évidemment... quand je dis mieux se faire connaître, c'est mieux faire connaître ses missions »

215 « Comme il n'y a pas de réelle maîtrise dans le groupe projet, ce genre de choses, et qu'on a besoin d'un produit qui sorte assez rapidement, on est parti sur un périmètre initial qui va recouvrir ce qui se fait actuellement, avec l'EAD, ce genre de choses, et on a une ouverture, à l'intérieur, pour le web sémantique. [...] Le prestataire, dans sa réponse, devra nous montrer, nous prouver, qu'il peut nous accompagner, justement, au passage au web sémantique [...] On ne savait pas avec quoi partir exactement. »

216 « Il y a tout un tas de couches supplémentaires mais qui pour moi en tout cas, ne sont pas matures, [...] par exemple tous les mécanismes d'inférence que tu vas avoir au niveau des schémas OWL [...] par contre, avec les quelques briques qui sont en bas, tu fais pas mal de trucs. Le gros du boulot, c'est de se mettre d'accord sur les vocabulaires et tout ça, c'est pas tant une question de technique qu'une question de travailler ensemble. »

Ces éléments coïncident avec les objectifs mis en avant dans la politique d'archivage du projet, qui rappelait effectivement que « le fait d'organiser la bonne conservation du capital informationnel immatériel lié aux missions de service public permet de développer la communication de l'institution. » Ainsi, c'est davantage l'enjeu de la mise en valeur du service et de la visibilité des documents conservés qui est réellement perçu par les archivistes, notamment par les directeurs et leurs adjoints, qui voient avant tout le potentiel du web des données à travers le prisme des partenariats, notamment européens et internationaux, qui peuvent être mis en place, particulièrement en ce qui concerne la participation à des portails.

Description archivistique et web sémantique : la question de la compatibilité des pratiques papier avec la culture de la donnée

Nous évoquons dans la première partie de notre analyse des résultats le fait que la manière de concevoir les instruments de recherche pouvait être bouleversée par les modifications engendrées par le numérique, et nos différents interlocuteurs soutenaient l'idée selon laquelle ils demeuraient pertinents et adaptés aux pratiques imposées par les normes de description archivistique. Pour autant, nous l'avons vu, le web sémantique, notamment parce qu'il repose sur le modèle du triplet RDF, suppose la répétition des différents éléments qui le composent et semble alors peu compatible avec des procédures de description documentaire qui sont fondées sur des règles d'interdépendance et d'héritage. Afin d'interroger la compatibilité de ces deux conceptions de la description de l'information, nous avons creusé ces questions avec les archivistes, qui doivent par ailleurs préparer le terrain pour l'arrivée du modèle Records in Context. Pour Delphine Jamet, ces éventuelles difficultés sont bel et bien perçues, puisqu'elle considère que « le problème c'est la hiérarchisation, on ne répète pas l'information, on la met au niveau supérieur, [...] c'est quand même une pratique bien ancrée donc c'est compliqué. » En effet, en termes d'interopérabilité entre la conception actuelle, classique, des instruments de recherche et les technologies du web sémantique, c'est avant tout la problématique de leur structuration, qui dépend fortement de règles d'héritage entre différents niveaux de description, qui peut constituer un obstacle. Pour autant, d'après Pascal Geneste, la manière dont ils sont constitués fait encore sens aujourd'hui, au regard des pratiques de description

archivistique, dans leur globalité : « Pour l'instant, on considère que sans instrument de recherche, une archive, ou une donnée, ne peut pas être retrouvée. Il y a une dispersion de la production archivistique qui peut nous être versée, qui n'est pas lisible si on ne la met pas en perspective de manière large. » Ceci étant, ce dernier émet l'idée d'une évolution de la définition, ou du moins une redéfinition, de la portée des instruments de recherche.²¹⁷

Les instruments de recherche subissent alors de plein fouet les mutations liées à l'environnement numérique, et s'ils sont toujours considérés comme des outils pertinents, voire fondamentaux, pour la plupart des archivistes interrogés, il s'agit effectivement de faire évoluer leur nature et de les concevoir d'une nouvelle manière, en prenant en compte les nouveaux usages qui peuvent en être faits et les enjeux naissants de structuration des données. L'émergence du modèle de description archivistique Records in Context et la conception de l'ontologie qui l'accompagne, permet effectivement d'interroger cette compatibilité. Delphine Jamet rappelle ainsi que les pratiques de description actuelles coïncident parfaitement avec les évolutions prévues par RiC :

« On fait déjà plus ou moins ça, [...] les éléments qu'on met, nous, par exemple déjà dans ISAD(G) ou dans l'EAD, c'est des éléments qu'on va retrouver dans RiC. On a bien renseigné zone par zone donc, du coup, finalement champ par champ de l'EAD, donc on va pouvoir redispacher ça correctement dans RiC, au final. C'est de segmenter tout comme ça, qui va [changer la donne]. »

En effet, d'après ses propos, c'est plutôt la manière de concevoir les différents éléments de la description archivistique qui va évoluer et il s'agirait alors de réfléchir à la manière dont ceux-ci vont pouvoir être intégrés pour donner du sens aux données au sein des interfaces de consultation.²¹⁸ Pour Amélie Antoine, la difficulté à se représenter une telle évolution semble davantage découler d'un manque de visibilité généralisé : « J'attends de voir un petit peu, j'attends de voir les évolutions concrètes. Je ne peux pas vraiment en parler. »

217 « La notion d'instrument de recherche peut être très large. Un instrument de recherche, ce n'est pas forcément un inventaire analytique qui va décrire, à la pièce, un document ou une donnée, ça peut être une description, au niveau haut, d'un ensemble de données, qui peuvent être très larges. Les instruments de recherche évoluent dans le temps et on est passé d'une pratique érudite, dans les outils de recherche comme dans le traitement archivistique, à, aujourd'hui, une gestion des flux ou des masses et une gestion plus globalisée des outils de recherche et plus diversifiée. »

218 : « Je pense que c'est plutôt une façon de réfléchir à la description, qui va changer, c'est-à-dire [...] que ça simplifie certaines choses. Ça nous oblige à nous poser plus de questions sur quelle description on veut mettre. »

En revanche, des attentes en termes de rationalisation des pratiques, puisque RiC est censé permettre de fédérer les pratiques découlant des normes de description existantes, sont clairement exprimées par nos interlocuteurs, qui considèrent que la multiplication des normes va à l'encontre d'une harmonisation de la description archivistique.²¹⁹ Ce point de vue, qui rejoint l'idée que le numérique aurait d'ores et déjà permis de restructurer les pratiques des archivistes autour d'un corpus normatif commun, est partagé par Pascal Geneste, puisque celui-ci admet que RiC « englobe finalement les quatre normes internationales qui coexistaient auparavant, et dont on usait assez peu dans leur globalité. »

Lorsque nous interrogeons Amélie Antoine au sujet de la compatibilité entre les pratiques des archivistes et les principes du web sémantique, celle-ci met en avant l'intérêt de la mise en relation des données :

« La fiche producteur peut déjà être en lien avec d'autres notices d'autorité, de différents autres organismes, donc que je pense que cette pratique peut totalement s'insérer. Tout est inter-relié, parce qu'on a des instruments de recherche électroniques [...] avec un export en XML-EAD et EAC pour les fiches producteur. »

L'enjeu de l'interconnexion des différents éléments permettant de traiter les archives est donc clairement perçu, dans la mesure où elle doit permettre de créer du sens à partir des outils et techniques déjà en place dans les services d'archives. C'est une vision qui est globalement partagée par les archivistes interrogés, Delphine Jamet rappelant que « ça paraît logique aussi, c'est très logique tous ces triplets, le fait que tout rebondisse avec le web, [...] le but c'est de faire des rebonds tout le temps, puisqu'on va forcément retrouver des choses ailleurs, donc moi je trouve que c'est totalement approprié. »

Pour autant, les répercussions sur la manière dont les publics vont pouvoir s'approprier les outils qui vont découler de cette nouvelle manière de concevoir la description archivistique ne sont pas encore très claires pour nos différents interlocuteurs, et il s'agit alors d'interroger les facteurs qui entravent cette visibilité. C'est l'exemple des

219 Ainsi, d'après Amélie Antoine : « Ce que j'attends c'est vraiment une évolution au niveau des descriptions, une harmonisation et une optimisation. C'est-à-dire qu'on aille vraiment plus loin, aujourd'hui je trouve que c'est compliqué, on a ISAD(G) d'un côté, on a ISAAR qui fait peur à plupart des archivistes qui ne sont pas des fous de la normalisation. Je pense qu'il faut une harmonisation entre les normes et un outil, enfin une norme, qui reprend l'ensemble des champs, notamment les champs obligatoires, des différentes normes, d'harmoniser leurs champs... »

instruments de recherche et de leur mise en ligne, comme nous le supposions, qui permet de faire la transition entre l'impact du web sémantique, les pratiques et les enjeux en termes de restitution auprès des publics. Ainsi, Delphine Jamet admet que « c'est la façon dont on publie ces instruments de recherche qui n'est pas forcément pertinente. »²²⁰ Dès lors, comme nous l'avions évoqué, si la rédaction d'instruments de recherche demeure indispensable pour nos interlocuteurs, pour qui elle est fortement ancrée dans le cœur de métier, elle est beaucoup moins évidente pour les publics, notamment pour les usagers des sites et portails d'accès aux documents conservés par les archives. En effet, les internautes en question sont systématiquement confrontés, encore actuellement, à cette manière de faire et de structurer l'information, lorsqu'ils doivent rechercher un document. Ainsi, comme le résume Agnès Vatican, « là où à mon sens, on a encore des progrès à faire, c'est effectivement dans les interfaces de consultation. »

III.3. L'enjeu de l'accès et de la diffusion : quel potentiel pour les publics ?

L'appropriation par les usagers : entre doute et manque de repères

À l'occasion d'une présentation de l'ensemble du projet aux RASAD en février 2017, l'équipe présente insistait sur le nouvel impératif introduit par la loi Lemaire sur la République Numérique du 7 octobre 2016, à savoir la création d'un « service public de la donnée » auquel toutes les administrations doivent participer. C'est donc sous cet angle qu'est mise en valeur la mise à disposition de données de référence, et donc le recours à un référentiel tel que proposé par l'équipe du projet SAEM. La question de l'ouverture des données, à travers la garantie de leur qualité ainsi que leur mise en ligne pour qu'elles puissent être consultées par les utilisateurs, est donc l'une des composantes centrales du projet SAEM. Le CCTP du module référentiel rappelle par ailleurs cette volonté, en présentant les enjeux de la mise en relation des informations à travers les technologies du web sémantique :

« Chaque document pris isolément n'est porteur que de peu d'information. L'enjeu est donc de pouvoir le rattacher à un corpus significatif pour permettre la compréhension de toutes ses dimensions. Le contexte de production a déjà été évoqué, mais il est également nécessaire de documenter son lieu de production, de rattacher son contenu

²²⁰ Elle précise par la suite : « Ça me semble assez pertinent pour les archivistes, par contre c'est la façon dont on les exploite effectivement, qui peut ne pas être pertinente ni pour les archivistes d'ailleurs, ni pour le public [et] éclater tout ça peut être effectivement une bonne solution. »

à des termes d'indexation qui permettent de le classer et de conserver sa dimension temporelle. Chacune de ces mises en relation augmente son potentiel de découverte par les utilisateurs et la mise en contexte de son contenu informationnel. »

La dernière phrase indique ici clairement que la préoccupation de l'accès et de la réutilisation par les utilisateurs n'est pas laissée de côté. Il s'agit d'un élément central, qui doit être pris en compte pour véritablement s'inscrire dans un web des données liées et ouvertes. De la même manière, dans une note à l'intention des présidents du Conseil départemental de la Gironde, l'équipe du SAEM girondin défend son projet en mettant l'accent sur la volonté de « compléter l'offre de service aux citoyens, actuellement proposée par la direction des archives départementales, pour leur permettre d'accéder, depuis chez eux ou en salle de lecture, aux contenus archivés librement communicables, mais également de constituer un socle de confiance pour bâtir l'administration publique de demain », avant de rappeler que la création d'un référentiel de données s'inscrit dans un « volet communication à destination des citoyens des informations publiques librement communicables. » Dès la rédaction de la politique du système d'archivage, en 2013, ces questions représentent l'un des enjeux du projet, et ce avant même le choix d'un recours à un référentiel de données. En effet, en plus des différentes entités concernées par l'archivage, telles qu'elles sont prévues par le modèle OAIS, la politique d'archivage intègre également, parmi les acteurs concernés par le SAEM girondin, les « usagers externes ». Toujours selon ce document, cette catégorie réunit alors « grand public, chercheurs [et] utilisateurs d'un service public dématérialisé. »

Lorsque nous abordons la question du potentiel du web sémantique pour la recherche dans le cadre de la prochaine version du site des Archives départementales, Damien Dudezert admet : « Alors, [l'intérêt] je le vois plus pour les publics [...] par rapport aux publics, je vois surtout une ouverture dans la recherche, c'est faciliter la recherche [...] ils vont lancer une recherche, ils vont pouvoir découvrir plein de sujets qui vont aller à côté, qui vont être connexes et qui vont permettre d'avoir quelque chose de bien ciblé et de construit. » Au-delà des questions technologiques et de l'indexation des documents, ce sont les questions des usages liés aux interfaces de consultation et celle du potentiel pour les usagers qui se posent, comme le soulignent les propos tenus par A. Vatican : « Mais moi ce qui m'interroge beaucoup c'est les usages, je pense que c'est

très important. » Pascal Geneste rappelle lui aussi l'objectif de la manœuvre : « le but c'est bien que l'utilisateur puisse connaître et puisse accéder à l'ensemble des documents que nous conservons. » C'est donc la question des stratégies de diffusion et des outils à mettre en œuvre qui ressort, dans la mesure où les services d'archives souhaitent évidemment augmenter la visibilité de leurs fonds et promouvoir l'accès aux documents et aux données. Pour autant, ce n'est souvent pas par ce biais-là que sont perçus les apports du web sémantique dans le cadre du SAEM, puisque comme nous l'avons vu, elles ont en grande partie été envisagées comme des outils de gouvernance et de simplification des versements. Dès lors, un manque de visibilité règne sur la façon dont les usagers vont pouvoir bénéficier des évolutions en termes de description archivistique.

En effet, du côté de Logilab, Sylvain Thénault semble avoir des difficultés à concevoir comment ces usages vont se matérialiser à partir du moment où la solution d'archivage gironde sera déployée : « Je n'ai pas de mal à le croire, enfin à envisager comment ça pourrait marcher, effectivement ça paraît logique. Mais aujourd'hui, je n'envisage pas cet usage-là, enfin je n'en ai pas connaissance. » Ainsi, malgré une implication importante dans les différents aspects du projet, la question du potentiel en termes de recherche et d'accès aux documents n'est que peu perçue, à ce stade, par le prestataire. Pourtant, c'est lui-même qui avait formé l'équipe du COPROJ aux technologies du web sémantique. Cela pose la question des raisons qui expliquent le recours aux technologies du web sémantique et nous permet alors de dresser le parallèle avec le billet *Linked Data Caution*²²¹ que nous citons dans notre première partie : les présentations, conférences et formations, et même plus largement les projets qui s'articulent autour du web des données ne sont-elles pas trop centrées sur l'aspect technologique plutôt que sur les questions liées aux usages qui peuvent en découler ? C'est en tout cas une interrogation qui permet de dresser le parallèle avec les doutes auxquels font face les archivistes quant à la viabilité de ces technologies.

Le retard par rapport à l'implémentation des outils et fonctionnalités propres au web social (indexation collaborative, espaces d'échanges, comptes personnalisés, interconnexion avec les réseaux sociaux, etc.) au sein des services d'archives peut

221 ROCHKIND, art. cit. p. 50

également constituer un facteur pour expliquer une certaine réticence, ou une difficulté, à envisager la transition vers le web des données, qui est souvent considéré comme « l'étape d'après ». Ce retard s'observe également au sein même des équipes, qui ne sont pas formées à ces aspects. C'est en tout cas ce que nous explique la directrice des Archives départementales de la Gironde :

« Après, nous on est même pas au web 2.0. C'est-à-dire qu'on a à la fois potentiellement l'aptitude à être pilote sur le web sémantique, avec des fortes expertises dans nos équipes [mais] les outils du web 2.0, les pratiques collaboratives, tout ça n'est pas implémenté sur notre site, et c'est encore quelque chose qui nécessite un accompagnement au changement auprès des équipes. »

Sans exemple probant ni réelle preuve de l'efficacité du web sémantique pour les données archivistiques, c'est également un sentiment de méfiance qui peut émerger, notamment par rapport aux modifications en profondeur des habitudes de classement et de navigation qu'il pourrait théoriquement engendrer. Les propos tenus par Agnès Vatican constituent un indicateur de ce scepticisme : « Je pense qu'il faut faire attention à ne pas penser que la technologie qui vient sera toujours la meilleure. » Elle appuie son propos à partir d'un exemple, celui du portail national FranceArchives :

« Il y a un exemple, en ce moment dans la profession, qui nous amuse, c'est le portail FranceArchives [...] : les professionnels des archives, qui avaient l'habitude, ne retrouvent pas les circulaires, les instructions de tri, etc. Et le SIAF vient d'envoyer un message assez long, une sorte d'aide pour nous expliquer comment retrouver les circulaires, qui jusque-là étaient accessibles avec un classement thématique qui était parfaitement opérant pour des professionnels. »

S'il faut nuancer cette observation en se posant la question de la représentativité du portail FranceArchives par rapport aux initiatives qui émergent de ce contexte, le bouleversement des habitudes de navigation que peuvent engendrer les portails qui reposent sur le web des données est une source d'inquiétude pour notre interlocutrice. En effet, s'il est déjà difficile à appréhender pour les professionnels, qu'en est-il alors pour les publics ? Par ailleurs, si cette problématique mise en avant par A. Vatican n'a pas été systématiquement soulevée par les autres archivistes avec lesquels nous nous sommes entretenus, elle symbolise les doutes auxquels fait face une partie de la profession par rapport à la création de telles plateformes. Notre interlocutrice met effectivement en opposition les pratiques de navigation que supposerait le web des données avec les habitudes de « navigation hiérarchisée » préexistantes. Elle pointe alors ce qu'elle considère être « les limites d'un portail où

l'entrée ne se fait que par métadonnées. »²²² Ainsi, d'après elle, « le dépouillement ne peut pas être remplacé par une description par métadonnée » et avec le web des données, il y aurait alors « toute une partie d'imprévu qui est moins certaine que quand on prend un inventaire et qu'on le lit de A à Z. » Ces éléments, beaucoup plus surprenants, renforcent l'opposition précédemment mise en lumière, laissant supposer une véritable méfiance par rapport aux modes de recherche que favorise le web des données, qui devraient quant à eux justement promouvoir la sérendipité et les multiples « rebonds » évoqués par nos autres interlocuteurs. Et effectivement, A. Vatican perçoit bel et bien cet aspect à travers la logique de rebond entre les mot-clés, concepts, personnes etc : « On passe à côté de certaines choses, même s'il y a toutes les possibilités de rebonds, etc, qui permettent d'aller d'une notion à l'autre, ça je le sais, je le pratique. » Pour autant, elle considère qu'en la matière, le web des données, dont on pourrait justement penser que c'est le principal atout, n'est pas à la hauteur du glanage d'informations dans des documents finis et linéaires. Cela rejoint l'idée selon laquelle la culture de l'imprimé et du document serait encore très prégnante dans les rangs des archivistes, et pose la question de la restitution de leur travail au sein même des interfaces de consultation.

Les interfaces de consultation des archives au cœur des mutations liées aux évolutions du web

Dans la mesure où le référentiel de données du SAEM girondin est pensé pour être relié au logiciel métier des archivistes, les données qui y sont intégrées vont pouvoir être réutilisées dans le cadre de la prochaine version du site des Archives départementales de la Gironde. Le fait d'interroger les procédures mises en œuvre pour la refonte du site des Archives départementales de la Gironde nous permet alors d'aborder des questions transversales qui mettent en exergue la nécessité d'avoir des interfaces de consultation qui correspondent à une nouvelle manière de concevoir la donnée archivistique. Pour Pascal Geneste, l'objectif de ce nouveau site web est avant

222 « L'accès au catalogue, avec des modalités de navigation et de recherche de différents niveaux, [...] des rubriques à disposition des professionnels, [...] une navigation hiérarchisée et pour les professionnels que nous sommes, on sait où trouver l'information qui nous intéresse. »

« La consultation d'instruments de recherche, on va dire classiques, permet aussi [...] la sérendipité [alors que] le web de données, je ne dis pas qu'il l'exclut, mais en tout cas il le rend moins probable. »

« La métadonnée, elle enferme la recherche, elle enferme autour d'un terme, autour d'une notion, beaucoup plus que le feuilletage, qu'il soit en ligne ou sur papier. »

tout « de ne plus avoir qu'un seul outil, qui intègre la description des fonds et leur consultation, mais aussi tout ce qui peut être mis en ligne et disponible à un internaute. » Il dresse alors le parallèle avec le module référentiel de données du SAEM girondin, en soulignant que l'objectif, in fine, est bien celui d'une consultation par l'utilisateur et d'une restitution fédérée de l'information, qui ne dépendrait pas alors de son lieu de stockage :

« Le site internet c'est la fenêtre de la structure que l'on représente et c'est les missions essentielles de cette structure qui doivent être mises en avant. Et comme je le disais tout à l'heure, en parlant du module référentiel, pour moi, la mission essentielle, c'est de permettre la consultation d'une donnée, d'une archive, quelle qu'elle soit, où qu'elle se trouve. »

S'il a été difficile d'évaluer la perception réelle qu'ont les archivistes des technologies du web sémantique, puisqu'ils n'y ont pas forcément été confrontés directement et sont naturellement réticents à l'idée de s'exprimer sur un sujet qu'ils ne maîtrisent pas, l'exemple du site internet nous permet de voir que les enjeux leur sont finalement familiers et fait ressortir des besoins auxquels les outils du web sémantique peuvent potentiellement répondre. En effet, ils considèrent avant tout le site comme une vitrine et c'est la question de l'accès à l'information qui y est primordiale. Ce sont notamment les fonctionnalités de recherche qui pourraient alors bénéficier du recours à une description basée sur un modèle de graphe, comme le reconnaît Delphine Jamet : « Pour la recherche à facettes ça facilite vraiment la donne, pour les usagers. Parce que quand tu regardes Bach d'Anaphore, c'est basé sur une ontologie. Et finalement tu n'as plus de formulaire de recherche avancée ou particulier, c'est de la recherche simple [...] donc je pense que pour le lecteur c'est plus clair. » D'après elle, le recours à une description multidimensionnelle échappant aux limitations imposées par les instruments de recherche permettrait d'avoir « une recherche très simple, et pouvoir filtrer après par facettes. » Elle dresse alors le bilan selon lequel les fonctionnalités de recherche actuelles posent problème, même auprès des archivistes.²²³

La diversité des orientations suivies dans le cadre de la conception de ces interfaces de consultation et de ces sites web, qui dépendent directement des feuilles de route des services d'archives et des offres proposées par les prestataires, nous permet d'observer l'existence d'une « double tension », pour reprendre le terme employé par Agnès

223 « On a plein de gens qui ne comprennent pas comment ça marche, même en interne. »

Vatican.²²⁴ En partant de l'exemple de GAEL (Gironde Archives En Ligne), l'interface de consultation des AD de la Gironde, elle rappelle les problèmes que pose une orientation métier de la conception de l'interface : « Mais on a une interface qui est très orientée métier, et [...] on avait un indice de satisfaction qui était en dessous de la moyenne des Archives départementales. » De fait, cette double tension nous permet d'observer cette difficulté pour les services d'archives à passer d'une logique orientée métier, avec des interfaces de consultation qui demeurent intimement liées à la structure des instruments de recherche et à la description à plusieurs niveaux propres aux fonds d'archives, à une approche qui serait davantage orientée usager, à travers la prise en compte de leurs habitudes de navigation et des standards du web. Les enquêtes de satisfaction, d'après les dires d'A. Vatican, mettent en lumière cette difficulté à s'approprier les interfaces et les outils de recherche encore souvent proposés par les services d'archives. Celle-ci se pose par ailleurs la question du positionnement des services d'archives par rapport à ces deux cultures que tout semble opposer : « On est entre ces deux cultures, ces deux visions. Et la question, c'est effectivement où on se situe, quel équilibre on trouve entre les attentes des uns et des autres. » L'expression de recherche d'un équilibre est ici intéressante, puisqu'elle laisse supposer que le fait de succomber entièrement aux besoins des publics et d'abandonner l'affichage des informations selon le modèle « classique » de l'instrument de recherche ne constituerait pas nécessairement une solution idéale. Pour autant, une chose est claire : les éléments propres à la pratique du métier d'archiviste et aux procédures en matière de classement et de traitement des documents sont bien souvent trop apparentes sur les interfaces de consultation et peuvent alors entraver l'accès aux informations recherchées par les usagers, qui ne devraient pas être confrontés à ces aspects au cours de leur navigation.²²⁵ C'est également le constat que dresse Pascal Geneste, qui va jusqu'à évoquer un « carcan »

224 « Ce qu'on a vu, ce qui est arrivé après, et notamment les sites Internet, mais c'était aussi le cas dans ces premiers logiciels de gestion : il y a les gens qui ont le point de vue du « tout métier » et qui vont défendre avant tout le « tout métier », et puis il y a des gens qui vont être plutôt sur la problématique de la réception par l'utilisateur, qui peut être l'archiviste ou qui peut être le public. On est sur cette double tension, avec des prestataires qui sont très orientés métier, d'autres qui sont très orientés publics, et à la fin ça fait des sites assez différents. »

225 D'après A. Vatican : « Je pense qu'il faut faire attention, et considérer une stratégie plus globale, qui est quand même l'accès aux fonds, ce qui est important c'est de donner accès à nos fonds, parce que le lecteur derrière, l'EAD, Records in Context, etc., il ne connaît pas, il n'a pas à connaître et il ne veut pas connaître, [...] ce qui l'intéresse, c'est de savoir ce qu'on a. »

qui serait un véritable obstacle à la recherche et à la réutilisation des contenus disponibles sur les sites d'archives :

« Il faut aussi que l'internaute qui ne connaît absolument pas le cadre de classement des archives et qui voudrait simplement avoir une information sur sa famille, sur sa commune, son entreprise, sa société ou sur tel sujet sur lequel il pourrait travailler, son club de sport etc, puisse en trouver, sans tenir compte de ce carcan, qui nous est imposé en termes de gestion, en quelque sorte, mais qui ne devrait pas l'être à quelqu'un d'extérieur. »

En outre, et au-delà de la manière dont les utilisateurs vont pouvoir accéder à l'information, c'est la question de la restitution du travail des archivistes qui interroge nos interlocuteurs : « Ici, on a des niveaux d'expertise très pointus mais en réalité, si vous regardez notre site Internet, ce n'est pas quelque chose qui est restitué au public, en tout cas pour l'instant. » D'après A. Vatican, il semble donc y avoir une forte dissonance entre les compétences techniques que développent les archivistes qui traitent les fonds et la manière dont l'information est restituée au public.²²⁶ Cela laisse supposer qu'il y a un potentiel mal exploité en termes d'accès et de médiation.²²⁷ Ce sont avant tout les fonctionnalités de recherche qui sont remises en cause, leur forme actuelle ne semblant pas correspondre aux attentes et habitudes des internautes. Ainsi, de nombreux doutes subsistent quant à la stratégie à adopter en matière de diffusion des données, dans la mesure où les interfaces de consultation héritent encore des pratiques de description et d'une logique de portail local, et donc de silo d'information, faisant alors encore l'objet d'un certain cloisonnement.

La mise en valeur de la donnée archivistique : quelles stratégies pour quels publics ?

Nous l'avions vu à travers l'exemple des instruments de recherche, c'est avant tout la question de leur exploitation et de l'exposition des données qui y sont contenues par le biais des sites web d'archives qui préoccupe les archivistes aujourd'hui. Dans cette perspective, et afin de concilier les besoins des différentes catégories de publics qui

226 « Parce qu'on peut avoir la description la plus conforme, à la pointe de tout ce qui est attendu en termes de métadonnées, mais si l'interface de restitution au public n'est pas opérante, ce travail-là n'est pas perçu par le public. »

227 « L'interfaçage avec les métadonnées qui sont produites en interne pour le public, à mon sens ici c'est quelque chose qui n'a pas fait l'objet d'une réflexion assez poussée, donc c'est un peu l'objet, l'enjeu même, de la refonte : comment est-ce qu'à partir de ce travail énorme qui a été fait, de rétro-conversion, de production d'instruments de recherche électroniques normalisés, comment est-ce qu'on donne à voir cette richesse. »

utilisent ces interfaces, du chercheur au généalogiste en passant par les curieux d'horizons divers, les archivistes s'interrogent sur l'éditorialisation des contenus et l'adaptation des modalités d'accès. Ainsi, Damien Dudezert, qui a participé à la rédaction du cahier des charges du prochain sites des Archives départementales émet la proposition suivante : « [Il faudrait] essayer d'arriver à séparer la partie professionnelle, avec tous les gens qui connaissent le langage archivistique, et l'utilisateur standard, qui a envie de faire une recherche sur l'état civil, pour qu'il ne soit pas perdu dans un cadre de classement, qu'il ne soit pas obligé de passer par un instrument de recherche. » Cette réponse reflète alors la volonté de prendre en compte les différences expériences des usagers et de faire évoluer la manière de présenter les fonctionnalités de recherche et d'accès aux documents numérisés. Delphine Jamet reconnaît elle aussi ces divergences en termes d'expérience utilisateur :

« Quand on est un généalogiste, il faut la recherche la plus simple possible et la visualisation la plus simple possible de ce qu'on veut, et avoir son résultat. Après, quand on est un historien, il faut avoir nos documents bien remis dans le contexte, parce qu'on ne fait pas la même analyse du document s'il n'est pas remis dans son contexte. »

Elle admet alors qu'une adaptation de l'interface pourrait permettre de répondre à cette problématique : « Le mieux ce serait qu'on se connecte avec un certain profil, et on a des interfaces particulières. »

Pour Amélie Antoine, le fait de prendre en compte les différentes typologies d'usagers directement dans l'interface du site semble compliqué. En revanche, faciliter le travail des archivistes, en permettant aux services producteurs d'interagir avec eux et d'avoir des outils accessibles en ligne, peut constituer un enjeu : « Non, moi je ne suis pas pour, pas spécialement. Il faut que le site internet soit une vraie plateforme de travail, aussi bien pour nous que pour les administrations, services versants et services producteurs, pour qu'il y ait un véritable échange. » Si Agnès Vatican reconnaît que l'utilisation de GAEL est « facile [pour] un docteur en Histoire », elle admet qu'un « usager qui est moins expert, il appelle au secours tôt ou tard. » Pour autant, elle met l'accent sur cette difficulté à proposer des expériences utilisateur sur mesure, en rappelant que, d'après elle, les attentes et les préférences en termes d'ergonomie des interfaces demeurent relativement similaires pour les différentes typologies de publics : « Je ne suis pas favorable au fait de mettre les gens dans des cases, [...] en

réalité, on voit bien que l'usager expert, quand l'interface est simple, il va aller vers cette ressource-là, lui aussi. » Il ne s'agit donc pas, de privilégier telle ou telle catégorie d'utilisateurs, mais plutôt de trouver un équilibre entre les besoins des uns et des autres. L'exemple de la recherche, et plus particulièrement des utilisateurs du monde universitaire, s'inscrit dans ce questionnement autour de la diversité des usages. Ainsi, lorsque nous évoquons le potentiel du web des données dans le contexte des Humanités numériques, Agnès Vatan s'interroge encore une fois sur la place des interfaces :

« La question que j'ai, c'est finalement : est-ce que le public des chercheurs universitaires est si différent que ça dans ses usages que du grand public ? Les universitaires aujourd'hui, vont beaucoup vers Gallica, qui est une interface assez simple de consultation, et elle n'est pas très complexe non plus pour le grand public. [...] on voit bien que ces enjeux d'éditorialisation des contenus, à la BnF, il y a une grande attention qui est portée à cela. »

À travers ce même exemple, c'est la question de la réception par les publics qui est soulevée par notre interlocutrice, puisque avec une mise en ligne massive au sein de sites ou portails dédiés, la question de la réception, et par extension celle du coût et de l'utilité de ces démarches, ne peuvent plus être éludées.²²⁸

Ces éléments posent alors la question de la fréquentation des services en ligne et de la réutilisation des données, qui pour l'instant, d'après elle, n'ont eu qu'un retentissement limité dans le monde de la recherche. En effet, en évoquant un projet de mise en ligne des textes de Pierre Mauriac par des chercheurs de l'Université de Bordeaux, elle signale : « ils faisaient quand même des constats sur qui utilise ces ressources mises en ligne, qui étaient assez mitigés [...] finalement c'est la question de l'utilité secondaire [...] ensuite tout ça, ça sert à qui ? » Nous le voyons, avant même d'envisager une exposition des données de l'archivage dans le web des données, nos interlocuteurs s'interrogent sur la manière dont les utilisateurs vont pouvoir bénéficier de ces plateformes, d'autant plus que ces projets peuvent être coûteux et complexes à mettre en œuvre. Par ailleurs, le contexte de recherche de performance au sein des collectivités, précédemment mis en avant par notre interlocutrice peut expliquer le fait que la mise en ligne ne doit pas être considéré comme un objectif en soi, et que les

228 « Ce qui est intéressant, c'est que les humanités digitales amènent les universitaires à se poser la question de la réception de ce travail-là, parce que là pour le coup il y a quand même des sommes engagées qui sont visibles [...] un universitaire qui veut mettre en ligne tel corpus etc, là les coûts ne sont plus du tout masqués. »

services d'archives, notamment, doivent de plus en plus se poser la question de la réutilisation pour mettre en valeur et faire financer leurs projets. C'est donc la question de l'utilité secondaire, de la réutilisation par les publics, qui doit revenir sur le devant de la scène.

Pour Amélie Antoine, c'est davantage le contexte de l'ouverture des données publiques qui devrait permettre au web des données de trouver un écho favorable auprès des services d'archives, puisque les impératifs liés à la qualité des données et à leur harmonisation se rejoignent à cette occasion : « L'esprit derrière, c'est quand même le partage d'information, il y a aussi les questions de l'open data aujourd'hui [...] c'est plutôt ces questions sur la réutilisation des données, où j'ai plutôt vu la mise en œuvre. » Elle admet que c'est par ce prisme que peut être envisagée la mise en correspondance des données issues de l'archivage avec les standards du web des données, en raison de l'envergure qu'a pris le mouvement de l'open data dans les services publics en France : « Au niveau de la direction de la culture, au niveau de la direction de la modernisation, on est plutôt sur des questions d'open data [...] là c'est une vraie révolution [...] c'est une autre façon d'anticiper et de penser. » L'archiviste aurait alors, dans le cadre de telles inflexions vers une ouverture des données, un véritable rôle à jouer : « L'archiviste doit aussi s'adapter à des questions telles que l'open data, et il faut totalement associer l'archiviste, on ne peut pas s'en passer. »

En effet, le potentiel du web des données est principalement perçu par les archivistes à travers les questions d'open data et de transparence, qui ont trouvé une place prépondérante au sein des impulsions des collectivités ces dernières années. La transition vers l'open data est alors davantage considérée comme une révolution que les concepts, encore flous pour les archivistes, de web sémantique et de Linked Data. Pour autant, comme le rappelle Agnès Vatican, la prise en compte des enjeux liés à l'ouverture des données publiques constitue « une couche de plus [...] un régime de plus, effectivement, d'accès aux données, bien sûr. » Pour les services d'archives, les règles en termes de communicabilité compliquent le rapport à l'open data et à l'exposition massive des données dans des portails nationaux. Les ramifications juridiques en termes d'accès constituent alors un facteur de complexité supplémentaire auquel les services d'archives vont devoir faire face, comme le souligne

également Pascal Geneste : « il y a aussi la problématique de la communication des archives publiques, qu'il ne faut pas négliger quand on développe des outils de recherche. » Il s'agit alors d'un facteur supplémentaire permettant d'expliquer ce scepticisme par rapport à une exposition dans le web des données.

Au-delà de la participation au mouvement de l'open data, et dans la perspective d'une réutilisation des données, il semble alors préférable de réfléchir à la manière dont les données archivistiques vont pouvoir être mises en valeur sur le web. Dès lors, il s'agirait, d'après ce que nous avons pu observer, d'envisager les systèmes qui reposent sur de telles technologies par le biais d'initiatives en lien avec des réalisations plus concrètes et attractives pour les publics, notamment en matière de data visualisation. C'est d'ailleurs cette manière de mettre en valeur la donnée qui est évoquée par Delphine Jamet, lorsqu'elle se réfère à un projet qui a retenu son attention : « Un projet américain, d'une université américaine, qui agrège les notices d'autorité et ils en font des graphes. » Agnès Vatican, lorsqu'elle évoque la problématique de la mise en valeur des données de la recherche, considère également que ce sont les dispositifs d'exposition des données qui doivent faire l'objet d'une attention particulière : « il y a des gens qui font des choses intéressantes, qui essaient de mixer l'audiovisuel, de combiner toutes ces ressources du web pour ouvrir leur recherche, la diffuser. » Dans cette perspective, plutôt que de mettre l'accent sur les technologies utilisées, il semble crucial de favoriser une réflexion concernant la manière d'exposer les données et d'envisager des dispositifs de restitution qui coïncident avec les évolutions du web et les attentes des internautes, afin de rendre possible une réelle valorisation, mission de la plus haute importance pour les services d'archives aujourd'hui, d'un patrimoine culturel. Ce sont alors des éléments qui feront certainement l'objet d'une réflexion plus poussée dans le cadre du projet girondin, à la suite de la mise en production des trois modules et à l'occasion de la création d'un entrepôt de métadonnées reposant sur l'ontologie RiC-O, interrogeable via un endpoint SPARQL²²⁹ et intégré à un nouveau module de pilotage²³⁰, mais aussi lorsque cette architecture devra être interconnectée avec le SIA des Archives départementales ainsi que leur futur site web.

229 SPARQL Protocol and RDF Query Language, le langage de requête du web des données.

230 Cf. planning et vision fonctionnelle cible du SAEM en annexe 17 et 18.

Conclusion

À travers cette étude, nous souhaitons interroger les différentes formes de potentialité que revêtent les technologies du web sémantique pour les services d'archives. En ce sens, l'analyse d'un système d'archivage électronique qui intègre au sein même de son architecture ces considérations, à défaut de nous permettre de nous positionner de manière définitive, laisse entrevoir les possibilités qui émergent à travers le recours à des dispositifs d'harmonisation des données. L'étude des procédures liées à la création d'un référentiel de données et les entretiens que nous avons pu mener nous permettent alors de comprendre en quoi les archivistes peuvent bénéficier de certaines briques technologiques du web sémantique, et plus particulièrement le recours à des identifiants uniques et pérennes et à des vocabulaires contrôlés en SKOS. Perçu avant tout comme un outil de gouvernance au service de différents acteurs de l'archivage, plus nombreux que jamais dans un environnement numérique, le référentiel de données du SAEM girondin a pour principal objectif la rationalisation des pratiques et la simplification du processus de gestion des versements. Si de tels outils sont encore considérés comme nouveaux par les archivistes, l'acculturation à la donnée et aux spécificités de l'archivage numérique semblent déjà, comme le soulignent les propos tenus par nos différents interlocuteurs, faire l'objet d'une réflexion et être intégrés au sein des feuilles de route des services d'archives.

Pour autant, ces mutations entraînent de nouvelles compétences et un repositionnement de l'archiviste, qui doit être à même de dialoguer avec les informaticiens, tout en mettant en valeur ses compétences propres. Cette conduite du changement ne concerne pas uniquement les aspects techniques mais questionne également l'impact de la culture professionnelle, notamment en termes de méthodes de gestion de projet. Le développement d'un vocabulaire commun et d'échanges équilibrés entre ces différents acteurs est alors le préalable nécessaire à un enrichissement des interactions interprofessionnelles et à la mise en œuvre d'une gouvernance de l'information.

Considérée comme un réel besoin plutôt que comme une contrainte, l'évolution des

pratiques de description archivistique vers un modèle unifié, pensé pour être interopérable avec les standards du web des données et dans l'optique d'une ouverture vers l'ensemble des acteurs du patrimoine et de la culture, devrait contribuer à cette conduite du changement progressive. Par ailleurs, si l'impact des technologies du web sémantique est encore flou pour les archivistes et qu'il a été difficile d'interroger directement ces aspects, les enjeux en termes d'ouverture et de simplification de la recherche sont bel et bien envisagés. De fait, si la rédaction d'instruments de recherche est encore profondément ancrée dans les pratiques de ces professionnels, ils reconnaissent la nécessité d'élargir leur manière de concevoir la description archivistique.

Nous avons également pu observer l'importance de l'implication des décideurs et des impulsions au niveau central, tant pour améliorer la visibilité des fonds et des institutions que pour assurer une montée en compétences des archivistes à travers des programmes de formation, qui pour l'instant ne semblent pas adéquats, et des solutions ouvertes et partagées. Le programme VITAM²³¹, dont l'offre est encore à l'étude par les équipes du projet SAEM, est symbolique de ces orientations, comme le soulignait le rapport de Christine Nougaret, que nous citons en introduction :

« Compte tenu des défis technologiques à relever, l'appui d'un centre de service national pour le numérique dans les archives est nécessaire, notamment pour assurer la pérennité des projets d'archivage numérique développés par l'État dans le cadre du programme VITAM. »²³²

Si notre protocole d'expérimentation ne nous permet pas de tirer de conclusion qui représente l'ensemble des mutations auxquelles sont soumis les services d'archives en France, en raison de la diversité des orientations suivies en matière d'archivage électronique, il a été propice à un questionnement sur la voie que devraient suivre les interfaces de consultation dans les temps à venir. D'après ce que nous avons pu observer, il s'agirait de trouver un équilibre entre les pratiques de description archivistique et les attentes des publics, qui n'ont pas à être confrontés aux spécificités du traitement des archives au cours de leur navigation.

Pour autant, nous l'avons vu, un manque de visibilité et une absence de points de

231 Valeurs immatérielles transmises aux archives pour mémoire.

232 NOUGARET, Christine. Une stratégie nationale pour la collecte et l'accès aux archives à l'ère numérique [en ligne]. 2017. [Consulté le 16 août 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/28204701>

repères concrets obstruent encore aujourd’hui l’ouverture des données des services d’archives, et notamment la manière dont les usagers vont pouvoir profiter de ces interconnexions, malgré des intentions clairement énoncées au niveau central. Le développement, au fil du temps, du portail FranceArchives, avec la participation accrue de l’ensemble des services d’archives français, et la mise en application du modèle Records in Context très prochainement, pourraient contribuer à donner une nouvelle dynamique à ces impulsions. Si nous supposons, dès notre introduction, une éventuelle corrélation entre ces évolutions et celles qui caractérisent le mouvement de l’ouverture des données publiques, nos entretiens nous ont réellement permis de prendre la mesure de cette interdépendance. En effet, le mouvement de l’open data et les objectifs de transparence et de réutilisation qui le caractérisent, semblent constituer un enjeu réellement stratégique pour les services d’archives publiques aujourd’hui. De fait, les Archives nationales viennent tout juste, le 21 juillet dernier, de publier un nouveau régime de réutilisation des instruments de recherche et des fonds qu’elles conservent, avec pour maîtres mots ouverture et gratuité.²³³ Il s’agit alors d’un véritable tournant pour la diffusion des archives, qui vont pouvoir bénéficier d’un rayonnement sans précédent. Au vu des ambitions du portail FranceArchives et des objectifs affichés par la feuille de route du ministère de la Culture concernant les métadonnées culturelles, il ne fait nul doute que l’archiviste a un rôle considérable à jouer dans cette nouvelle manière de penser l’accès aux archives, comme le rappelait déjà Françoise Banat-Berger en 2014 :

« Avec leur savoir-faire historique et leurs outils conceptuels renouvelés, les archivistes sont à même d’identifier les données de référence, archives de demain, et de promouvoir leur archivage pérenne. Il importe donc qu’ils participent à la réflexion et fassent entendre leur voix singulière dans le débat actuel sur les sources de la connaissance en devenir. »²³⁴

Dès lors, l’émergence d’outils reposant sur les technologies du web sémantique, visant à assurer l’interopérabilité et la qualité des données, dans l’optique de leur ouverture, devrait permettre aux services d’archives de trouver une place au sein de ces dispositifs et d’assurer la mise en valeur de la richesse des fonds qu’elles conservent.

233 Réutilisation des informations publiques. *Archives nationales (France)* [en ligne]. [Consulté le 25 août 2017]. Disponible à l’adresse : <http://www.archives-nationales.culture.gouv.fr/reutilisation-des-donnees-publiques>

234 BANAT-BERGER, Françoise et NOUGARET, Christine. Faut-il garder le terme archives ? Des « archives » aux « données ». *La Gazette des archives*. 2014. Vol. 233, n° 1, pp. 7-18. DOI 10.3406/gazar.2014.5121.

Table des matières

Remerciements.....	3
Table des abréviations.....	4
Sommaire.....	5
Introduction.....	6
Première partie : l'impact du numérique et des évolutions du web sur le traitement des données archivistiques.....	10
I. L'impact du numérique sur la pratique de l'archivage.....	10
I.1. Les archives, produit documentaire de l'activité humaine.....	10
Le document d'archives, éléments de définition.....	10
Les fondements de l'archivage : les principes de respect du fonds et la description à plusieurs niveaux.....	12
Cycle de vie et processus de collecte, quelle application dans l'environnement numérique ?.....	13
I.2. De ISAD(G) à EAD : l'évolution des normes de description archivistique....	15
Le quatuor normatif de la description archivistique.....	15
Description archivistique et métadonnées, un enchevêtrement naturel ?.....	17
I.3. Du document à la donnée, quelles implications pour les instruments de recherche ?.....	19
Document, ressource ou donnée : l'enjeu de la granularité.....	19
L'instrument de recherche, document ou ensemble de données ?.....	21
II. De l'élaboration de modèles conceptuels à la mise en place de référentiels : les enjeux d'une gouvernance de l'information et des données.....	25
II.1. La gestion de l'information à l'ère du numérique, une fonction transversale	25
La gouvernance, un concept défini par son caractère systémique.....	25
Les apports des compétences de l'archiviste à cette vision globale.....	26
La collaboration avec les directions informatiques, l'occasion d'une réévaluation du positionnement de l'archiviste.....	28
II.2. L'interopérabilité : condition première de la gouvernance des données.....	29
Les outils de la gouvernance : de la politique d'archivage au référentiel de données.....	29
Modèles conceptuels et standards d'échanges : l'archivage électronique dans un contexte de gouvernance.....	32
III. Optimiser la gouvernance dans un service d'archives : les potentialités des technologies du web sémantique.....	35
III.1. Les fondements du web sémantique : créer du lien et donner du sens aux données.....	35
Le web sémantique, une vision présente dès les prémices du web.....	35
Ontologies et vocabulaires : le socle de la sémantisation des données.....	38

Web des données et techniques documentaires, quelle imbrication ?.....	39
III.2. Web des données et archivage : quel potentiel pour la protection et la valorisation d'un patrimoine.....	40
Le web des données, un enjeu perçu à l'échelle de l'ensemble du secteur culturel.....	40
L'application au domaine des archives : spécificités et perspectives.....	42
« Records in Context », un modèle conceptuel et une ontologie dédiées à la description archivistique.....	45
III.3. Le web des données, entre enjeux de gouvernance et stratégie de diffusion	47
Description archivistique et données liées, entre antagonisme et complémentarité.....	47
Le recours aux technologies du web sémantique au service de la gouvernance de l'information ?.....	48
La diffusion des données archivistiques sur le web, quelle réalité pour les usagers ?.....	49
Deuxième partie : l'élaboration du protocole d'expérimentation.....	52
I. L'étude de cas à partir de l'exemple du SAEM girondin : attentes et objectifs.....	52
I.1. Le choix de l'étude des procédures du SAEM : le contexte de l'expérimentation.....	53
Contexte et considérations pratiques de l'expérimentation.....	53
Un questionnaire initial : favoriser l'approche comparative ou l'étude de cas ?.....	53
I.2. Le système d'archivage électronique mutualisé girondin, un exemple représentatif ?.....	54
Un projet considéré comme pionnier dans le domaine de l'archivage électronique pour les collectivités.....	54
Le module référentiel : pierre angulaire de l'harmonisation des données et symbole du recours aux technologies du web sémantique.....	57
Un mode de gouvernance particulier : l'organisation du comité projet et l'impact de la mutualisation.....	58
I.3. L'approche suivie et ses différentes étapes.....	59
L'analyse de la production documentaire et des procédures du projet.....	59
La manipulation des outils et l'étude des tickets et user stories.....	60
II. L'approche qualitative pour compléter et approfondir les résultats.....	61
II.1. Le choix de l'entretien semi-directif.....	61
Un objectif : s'intéresser à la perception et au ressenti des archivistes et dépasser la vision « objective » des procédures.....	61
II.2. La réalisation du guide d'entretien.....	62
Le choix des questions et l'articulation avec les hypothèses.....	62
Un outil évolutif et prenant en compte de la pluralité des types d'acteurs interrogés.....	63
II.3. Définition du panel et déroulement des entretiens.....	64
Le choix des personnes interrogées : critères de sélection et objectifs ciblés...64	
Prise de contact et modalités des entretiens.....	66

III. Procédure d'analyse et d'interprétation des résultats.....	68
La mise en regard des données issues de l'étude des procédures avec les résultats des entretiens.....	68
Troisième partie : résultats de l'expérimentation.....	69
I. La transition numérique des services d'archives, symbole d'une bascule dans les pratiques professionnelles.....	69
I.1. L'archivage électronique, signe avant-coureur d'une montée en compétences inéluctable ?.....	69
Une incompatibilité entre le rythme des évolutions technologiques et le besoin de stabilité des services d'archives ?.....	73
La normalisation, garante de l'interopérabilité entre les pratiques de l'archivage papier et les apports du numérique.....	74
I.2. Des besoins en termes d'acculturation et de conduite du changement.....	78
L'appropriation des outils de l'archivage électronique par les archivistes : l'enjeu de la conduite du changement.....	78
Du document à la donnée : quelle place pour l'instrument de recherche ?.....	80
Des carences en termes de formation et de sensibilisation à la donnée et à l'archivage numérique ?.....	81
II. Le nouveau paradigme de la gouvernance de l'information : vers davantage de transversalité.....	84
II.1. L'essor de la collaboration interprofessionnelle, source de difficulté ou facteur d'enrichissement ?.....	84
II.2. Gouvernance et gestion de projet, quelle imbrication et quelles conséquences pour l'archiviste ?.....	89
Méthodologie de gestion de projet et rôle des décideurs : quel impact sur la gouvernance ?.....	89
Du référent au spécialiste : la question du positionnement d'acteurs clé au sein même des services.....	91
III. Web sémantique et données archivistiques, un enchevêtrement souhaité et souhaitable ?.....	93
III.1. Harmonisation des données, automatisation du traitement des documents : le web sémantique au service des archivistes.....	93
Le module référentiel du SAEM : un outil de gouvernance au service de ses utilisateurs.....	93
Le recours à des identifiants uniques et pérennes.....	96
Vocabulaires contrôlés et ontologies, désambiguïser les données pour favoriser l'interopérabilité.....	99
Les notices d'autorité et les profils SEDA pour encourager la réutilisation et faciliter les échanges.....	101
III.2. Les technologies du web sémantique, quel impact sur les pratiques des archivistes ?.....	102
Des mutations technologiques caractérisées par un manque de visibilité.....	102
Description archivistique et web sémantique : la question de la compatibilité des pratiques papier avec la culture de la donnée.....	105

III.3. L'enjeu de l'accès et de la diffusion : quel potentiel pour les publics ?.....	108
L'appropriation par les usagers : entre doute et manque de repères.....	108
Les interfaces de consultation des archives au cœur des mutations liées aux évolutions du web.....	112
La mise en valeur de la donnée archivistique : quelles stratégies pour quels publics ?.....	115
Conclusion.....	120
Table des matières.....	123
Bibliographie.....	127

Bibliographie

Archivistique et numérique : évolution des normes de description et des instruments de recherche

- ANGJELI, Anila. Bibliothèques, archives, identifiants : les enjeux de la normalisation. *Hermès, La Revue*. 25 novembre 2013. N° 66, pp. 167-172.
- BACHIMONT, Bruno. Ingénierie des connaissances et des contenus : le numérique entre ontologies et documents. *Hermès Science Publ.*, 2007. ISBN 978-2-7462-1369-2.
- BANAT-BERGER, Françoise et NOUGARET, Christine. Faut-il garder le terme archives ? Des « archives » aux « données ». *La Gazette des archives*. 2014. Vol. 233, n° 1, pp. 7-18. DOI 10.3406/gazar.2014.5121.
- BANAT-BERGER, Françoise. Les archives et la révolution numérique. *Le Débat*. 1 mars 2011. N° 158, pp. 70-82.
- BANAT-BERGER, Françoise. « Un métier à part entière, l'archiviste un généraliste de l'information » : qu'en est-il en 2012 dans le nouvel environnement numérique ? *La Gazette des archives*. 2012. Vol. 226, n° 2, pp. 117-126. DOI 10.3406/gazar.2012.4901.
- BANAT-BERGER, Françoise. De l'écrit à internet : comment archive-t-on l'immatériel ? *Pouvoirs*. 20 avril 2015. N° 153, pp. 109-124.
- BROCHIER, André. Instruments de recherche et Internet : référentiels et métadonnées. *La Gazette des archives*. 2010. Vol. 220, n° 4, pp. 113-125. DOI 10.3406/gazar.2010.4745.
- CAVALIÉ, Étienne. L'EAD, c'est pour l'import-export. *Bibliothèques [reloaded]* [en ligne]. 7 octobre 2013. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://bibliotheques.wordpress.com/2013/10/07/lead-cest-pour-limport-export/>
- CAVALIÉ, Étienne. Calames IT ? *Bibliothèques [reloaded]* [en ligne]. 21 juillet

2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://bibliotheques.wordpress.com/2014/07/21/calames-it/>
- CHABIN, Marie-Anne. *Le management de l'archive*. Hermès science, 2000. ISBN 978-2-7462-0107-1.
 - CHABIN, Marie-Anne. Document trace et document source. La technologie numérique change-t-elle la notion de document ? *Revue I3 – Information Interaction Intelligence* [en ligne]. 5 juillet 2004. Vol. 4, n° 1. [Consulté le 21 avril 2017]. Disponible à l'adresse : https://archivesic.ccsd.cnrs.fr/sic_00001020/document
 - CHABIN, Marie-Anne. Qu'est-ce qu'un document d'archives ? *TRANSARCHIVISTIQUE* [en ligne]. août 2017. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://transarchivistique.fr/quest-ce-quun-document-darchives/>
 - CŒURÉ, Sophie et DUCLERT, Vincent. Introduction / Une histoire à écrire. *Repères*. 11 mai 2011. pp. 3-8.
 - COLOMBANI, Louis. Instrument de recherche, données et document. *Anaphore* [en ligne]. 21 janvier 2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/instrument-de-recherche-donnees-et-document/>
 - COLOMBANI, Louis. La description archivistique à l'ère du numérique – Part 1. *Anaphore* [en ligne]. 3 juillet 2014. [Consulté le 17 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/la-description-archivistique-a-lere-du-numerique-part-1/>
 - COLOMBANI, Louis. La description archivistique à l'ère du numérique – Part 2. *Anaphore* [en ligne]. 8 juillet 2014. [Consulté le 21 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/la-description-archivistique-a-lere-du-numerique-part-2/>
 - DHÉRENT, Catherine, GRIMOÛARD, Claire Sibille De, OTT, Florence et BROULAND, Myriel. Comprendre les enjeux et les fondamentaux. *Documentaliste-Sciences de l'Information*. 5 juin 2009. Vol. 46, n° 2, pp. 32-42.

- FILIPPOZZI, Lorraine. Le web comme outil de diffusion des archives. *Ressi* [en ligne]. avril 2008. [Consulté le 17 avril 2017]. Disponible à l'adresse : http://www.ressi.ch/num07/article_044
- GALLAND, Bruno. Chapitre III. Les missions des archivistes. *Que sais-je ?* 27 décembre 2016. pp. 85-120.
- GALLAND, Bruno. Les archives. *Que sais-je ?* 27 décembre 2016. pp. 3-6.
- GRIMOÛARD, Claire Sibille-De. Avantages et inconvénients d'un encodage en EAD. *La Gazette des archives*. 2010. Vol. 220, n° 4, pp. 179-194. DOI 10.3406/gazar.2010.4757.
- GRIMOÛARD, Claire Sibille-de. Élaborer des normes de description... et les confronter à la pratique d'aujourd'hui. *La Gazette des archives*. 2012. Vol. 226, n° 2, pp. 165-177. DOI 10.3406/gazar.2012.4906.
- GRIMOÛARD, Claire Sibille-de. Les normes internationales de description archivistique : origines, développements, perspectives. *La Gazette des archives*. 2012. Vol. 228, n° 4, pp. 77-90. DOI 10.3406/gazar.2012.4985.
- GUYON, Céline. La pratique archivistique publique en France, entre adaptation et négociation. *Les Cahiers du numérique*. 17 juin 2015. Vol. 11, n° 2, pp. 77-114.
- LEMAY, Yvon et KLEIN, Anne. La diffusion des archives ou les 12 travaux des archivistes à l'ère du numérique. *Les Cahiers du numérique*. 9 janvier 2013. Vol. 8, n° 3, pp. 15-48.
- NAUD, Gérard. Un métier de notre temps : l'archivage. *La Gazette des archives*. 2012. Vol. 226, n° 2, pp. 43-60. DOI 10.3406/gazar.2012.4896.
- NOUGARET, Christine. *Une stratégie nationale pour la collecte et l'accès aux archives à l'ère numérique* [en ligne]. 2017. [Consulté le 16 août 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/28204701>
- PÉDAUQUE, Roger T. *Le Document à la lumière du numérique : forme, texte, médium : comprendre le rôle du document numérique dans l'émergence d'une nouvelle modernité*. C & F Éditions, 2006. ISBN 978-2-915825-11-4.

- ROUSSEL, Stéphanie. Le champ normatif de l'archivage électronique. *La Gazette des archives*. 2012. Vol. 228, n° 4, pp. 59-76. DOI 10.3406/gazar.2012.4984.
- ROUSSEL, Stéphanie. Congrès 2015 de la SAA. Episode 1 : une gouvernance unifiée pour les standards d'encodage archivistiques. *Modernisation et archives* [en ligne]. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://siaf.hypotheses.org/490>
- SALAÛN, Jean-Michel. La redocumentarisation, un défi pour les sciences de l'information. *Études de communication. langages, information, médiations*. 1 octobre 2007. N° 30, pp. 13-23. DOI 10.4000/edc.428.
- SALAÛN, Jean-Michel. Pourquoi le document importe. *INA Expert* [en ligne]. juin 2012. [Consulté le 26 août 2017]. Disponible à l'adresse : <https://www.ina-expert.com/e-dossier-de-l-audiovisuel-sciences-humaines-et-sociales-et-patrimoine-numerique/pourquoi-le-document-importe.html>
- SIBILLE, Claire. Les instruments de recherche électroniques. *La Gazette des archives*. 2006. Vol. 204, n° 4, pp. 131-150. DOI 10.3406/gazar.2006.3831.
- VERNUS, Pierre. Les instruments de recherche sont-ils adaptés à la demande des publics ? Le cas du public universitaire. *La Gazette des archives*. 2010. Vol. 220, n° 4, pp. 133-141. DOI 10.3406/gazar.2010.4747.
- *Article L211-1* [en ligne]. [Consulté le 3 mai 2017]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/affichCodeArticle.do?cidTexte=LEGITEXT000006074236&idArticle=LEGIARTI000006845559&dateTexte=&categorieLien=cid>
- International Council on Archives. *International Council on Archives* [en ligne]. [Consulté le 3 mai 2017]. Disponible à l'adresse : <http://www.ica.org/fr>

Gouvernance de l'information, archivage électronique et e-administration

- BANAT-BERGER, Françoise. « Un métier à part entière, l'archiviste un généraliste de l'information » : qu'en est-il en 2012 dans le nouvel environnement numérique ? *La Gazette des archives*. 2012. Vol. 226, n° 2, pp. 117-126. DOI 10.3406/gazar.2012.4901.

- BROCHIER, André. Instruments de recherche et Internet : référentiels et métadonnées. *La Gazette des archives*. 2010. Vol. 220, n° 4, pp. 113-125. DOI 10.3406/gazar.2010.4745.
- CENTRE INFORMATIQUE NATIONAL DE L'ENSEIGNEMENT SUPÉRIEUR. Le modèle de référence : l'OAIS. *CINES* [en ligne]. décembre 2016. [Consulté le 17 avril 2017]. Disponible à l'adresse : <https://www.cines.fr/archivage/un-concept-des-problematiques/le-modele-de-reference-loais/>
- COTTIN, Michel. Prendre des décisions fondées sur des preuves. *I2D – Information, données & documents*. 6 janvier 2017. Vol. 53, n° 4, pp. 41-42.
- COUCHET, Pierre. Le modèle de référence OAIS. *archivEngines* [en ligne]. 7 décembre 2011. [Consulté le 3 mai 2017]. Disponible à l'adresse : <https://archivengines.wordpress.com/2011/12/07/le-modele-de-reference-oais/>
- DEMOULIN, Marie et VERNUSSET, Amélie. La réversibilité des données et l'archivage électronique. *Les Cahiers du numérique*. 17 juin 2015. Vol. 11, n° 2, pp. 115-148.
- DHÉRENT, Catherine, GRIMOÛARD, Claire Sibille De, OTT, Florence et BROULAND, Myriel. Comprendre les enjeux et les fondamentaux. *Documentaliste-Sciences de l'Information*. 5 juin 2009. Vol. 46, n° 2, pp. 32-42.
- FRANCEARCHIVES – PORTAIL NATIONAL DES ARCHIVES. Archives électroniques | Normes et standards | SEDA. *FranceArchives* [en ligne]. 9 avril 2015. [Consulté le 3 mai 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/26287460>.
- GUYOT, Brigitte, FUZEAU, Pierre, COTTIN, Michel et RIETSCH, Jean-Marc. Repères conceptuels. *Documentaliste-Sciences de l'Information*. 15 avril 2013. Vol. 50, n° 1, pp. 26-37.
- HAGMANN, Jürg et ZELLER, Jean-Daniel. Gouvernance de l'information : véritable innovation dans la gestion de l'information ? *Les Cahiers du*

numérique. 17 juin 2015. Vol. 11, n° 2, pp. 15-36.

- JULES, Arnaud et LEBIGRE, Loïc. Information : le temps de la gouvernance. *Documentaliste-Sciences de l'Information*. 15 avril 2013. Vol. 50, n° 1, pp. 24-25.
- JULES, Arnaud et LEBIGRE, Loïc. Sources et ressources. *Documentaliste-Sciences de l'Information*. 15 avril 2013. Vol. 50, n° 1, pp. 60-61.
- JULES, Arnaud, MOURAIN, Jean, BLANGER, Jean-Pierre, MASSE, Claudine et CHABIN, Marie-Anne. Pratiques. *Documentaliste-Sciences de l'Information*. 15 avril 2013. Vol. 50, n° 1, pp. 50-59.
- JULES, Arnaud. Une politique de gestion des documents d'activité pour une gouvernance documentaire stratégique. *La Gazette des archives*. 2012. Vol. 228, n° 4, pp. 153-171. DOI 10.3406/gazar.2012.4991.
- KERN, Gilliane, HOLGADO, Sandra et COTTIN, Michel. Cinquante nuances de cycle de vie. *Les Cahiers du numérique*. 17 juin 2015. Vol. 11, n° 2, pp. 37-76.
- PERREIN, Jean-Pascal. Politiques et référentiels de la gouvernance de l'information. *GouvInfo IAI* [en ligne]. 29 mai 2012. [Consulté le 10 août 2017]. Disponible à l'adresse : <http://www.gouvinfo.org/IAI/politiques-et-referentiels-de-la-gouvernance-de-linformation/>
- PERREIN, Jean-Pascal. Les fondamentaux de la gouvernance de l'information : ce qui se fait en 2012. *GouvInfo IAI* [en ligne]. 5 juin 2012. [Consulté le 14 août 2017]. Disponible à l'adresse : <http://www.gouvinfo.org/IAI/les-fondamentaux-de-la-gouvernance-de-linformation-ce-qui-se-fait-en-2012/>
- ROUSSEL, Stéphanie. Le champ normatif de l'archivage électronique. *La Gazette des archives*. 2012. Vol. 228, n° 4, pp. 59-76. DOI 10.3406/gazar.2012.4984.
- SECRÉTARIAT GÉNÉRAL POUR LA MODERNISATION DE L'ACTION PUBLIQUE. Archivage numérique | Les documents de référence du S.I. de l'État. *Le portail de la modernisation de l'action publique* [en ligne]. [Consulté le 14 août 2017]. Disponible à l'adresse : <http://references.modernisation.gouv.fr/archivage-numerique>

- SERDA. *La Gouvernance de l'information numérique dans les organisations* [en ligne]. 2017. [Consulté le 11 juin 2017]. Disponible à l'adresse : <http://www.serdalab.com/Medias/Livres%20blancs-%c3%a9tudes%20gratuites/Le%205e%20rapport%20gouvernance%20Serda%202016.pdf>

Technologies du web sémantique et enjeux du web des données

- ANGJELI, Anila. Bibliothèques, archives, identifiants : les enjeux de la normalisation. *Hermès, La Revue*. 25 novembre 2013. N° 66, pp. 167-172.
- ANGJELI, Anila. Congrès 2015 de la SAA. Episode 2 : les Archives dans leur contexte (Records in context, RiC). *Modernisation et archives* [en ligne]. 5 novembre 2015. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://siaf.hypotheses.org/493>
- BACHIMONT, Bruno, GANDON, Fabien, POUPEAU, Gautier, VATANT, Bernard, TRONCY, Raphaël, POUYLLAU, Stéphane, MARTINEZ, Ruth, BATTISTI, Michèle et ZACKLAD, Manuel. Enjeux et technologies : des données au sens. *Documentaliste-Sciences de l'Information*. 31 janvier 2012. Vol. 48, n° 4, pp. 24-41.
- BERNERS-LEE, Tim. Semantic Web roadmap. *World Wide Web Consortium* [en ligne]. 14 octobre 1998. [Consulté le 3 mai 2017]. Disponible à l'adresse : <https://www.w3.org/DesignIssues/Semantic.html>
- BERNERS-LEE, Tim. Linked Data – Design Issues. *World Wide Web Consortium* [en ligne]. 27 juillet 2006. [Consulté le 3 mai 2017]. Disponible à l'adresse : <https://www.w3.org/DesignIssues/LinkedData.html>
- BRIATTE, Katell. Hadoc, un programme pour harmoniser les données culturelles. *Documentaliste-Sciences de l'Information*. 4 juillet 2014. Vol. 51, n° 2, pp. 56-67.
- CHARLET, Jean et KEMBELLEC, Gérald. Du web sémantique au web des données, quels enjeux professionnels ? *I2D – Information, données & documents*. février 2016. Vol. me 53, n° 2, pp. 54-55.
- COLOMBANI, Louis. Modèle conceptuel et description archivistique ? *Anaphore* [en ligne]. 10 février 2015. [Consulté le 17 avril 2017]. Disponible à l'adresse :

<http://www.anaphore.eu/modele-conceptuel-et-archives/>

- DALBIN, Sylvie, BERMÈS, Emmanuelle, ISAAC, Antoine, WENZ, Romain, NICOLAS, Yann, MERABTI, Tayeb, ANGJELI, Anila, FRANCCART, Thomas, ROZAT, Lise, VANDENBUSSCHE, Pierre-Yves, VATANT, Bernard, RAIMOND, Yves et COTTE, Dominique. *Approches documentaires : priorité aux contenus. Documentaliste-Sciences de l'Information*. 31 janvier 2012. Vol. 48, n° 4, pp. 42-59.
- DUHAMEL, Benjamin. *Les technologies du web sémantique et du record linkage au service de data.bnf.fr et du Linked Open Data culturel : étude sur les nouveaux paradigmes informationnels* [en ligne]. other. Université Lille 3, UFR DECCID, département SID, 2014. [Consulté le 17 avril 2017]. Disponible à l'adresse : https://memic.ccsd.cnrs.fr/mem_01081739/document
- EGAD. Records in Context – Modèle conceptuel | International Council on Archives. [en ligne]. [Consulté le 13 août 2017]. Disponible à l'adresse : <https://www.ica.org/fr/egad-ric>
- FILIPPOZZI, Lorraine. Le web comme outil de diffusion des archives. *Ressi* [en ligne]. avril 2008. [Consulté le 17 avril 2017]. Disponible à l'adresse : http://www.ressi.ch/num07/article_044
- FRANCCART, Thomas. Penser, modéliser (pour le web de données) – Part 1/2. *Anaphore* [en ligne]. 2 juin 2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/penser-modeliser-pour-le-web-de-donnees-part-1/>
- FRANCCART, Thomas. Penser, modéliser (pour le web de données) – Part 2/2. *Anaphore* [en ligne]. 4 juin 2014. [Consulté le 19 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/penser-modeliser-pour-le-web-de-donnees-part-2/>
- GRACY, Karen F. Archival description and linked data : a preliminary study of opportunities and implementation challenges. *Archival Science*. 1 septembre 2015. Vol. 15, n° 3, pp. 239-294. DOI 10.1007/s10502-014-9216-2.

- GRIMOÛARD, Claire Sibille de. Une évolution dans les pratiques descriptives – Vers un modèle conceptuel international pour la description archivistique. *Arbido*. 14 juin 2012. Vol. 1, n° 2, pp. 4-8.
- ILLIEN, Gildas. Le web sémantique, nouveau levier de la valeur pour les services d'information ? *I2D – Information, données & documents*. 6 janvier 2016. Vol. me 53, n° 4, pp. 59-60.
- INTERPARES TRUST. InterPARES Trust responds to EGAD-RiC. *InterPARES Trust* [en ligne]. 11 décembre 2016. [Consulté le 17 avril 2017]. Disponible à l'adresse : <https://interparestrust.com/2016/12/11/interparestrust-responds-to-egad-ric/>
- JONES, Mike. Comments on Records in Contexts. *Context Junky* [en ligne]. 1 février 2017. [Consulté le 1 mai 2017]. Disponible à l'adresse : <http://www.mikejonesonline.com/contextjunky/2017/02/02/comments-on-records-in-contexts/>
- MAIGNIEN, Yannick. Les enjeux du web sémantique. <http://www.parcoursnumeriques-pum.ca> [en ligne]. 27 mai 2015. [Consulté le 21 avril 2017]. Disponible à l'adresse : <http://parcoursnumeriques-pum.ca/les-enjeux-du-web-semantique>
- MENON, Bruno. Comprendre les standards du web de données. *I2D – Information, données & documents*. février 2016. Vol. 53, n° 2, pp. 32-34.
- MINISTÈRE DE LA CULTURE ET DE LA COMMUNICATION. Feuille de route – Métadonnées culturelles et transition Web 3.0 (janvier 2014) – Page 1. [en ligne]. [Consulté le 19 avril 2017]. Disponible à l'adresse : http://www.culturecommunication.gouv.fr/var/culture/storage/pub/feuille_de_route__metadonnees_culturelles_et_transition_web_3_0_janvier_2014/index.htm
- MINISTÈRE DE LA CULTURE ET DE LA COMMUNICATION. L'application GINCO – Ministère de la Culture. [en ligne]. [Consulté le 7 août 2017]. Disponible à l'adresse : <http://www.culturecommunication.gouv.fr/Divers/Harmonisation-des-donnees-culturelles/Referentiels/Les-vocabulaires-scientifiques-et->

techniques/L-application-GINCO

- NOUGARET, Christine. Une stratégie nationale pour la collecte et l'accès aux archives à l'ère numérique. *FranceArchives* [en ligne]. [Consulté le 31 juillet 2017]. Disponible à l'adresse : <https://francearchives.fr/fr/article/28204701>
- PAUILLAC, Myriam. Web des données et archives – quel intérêt ? *Anaphore* [en ligne]. 23 novembre 2014. [Consulté le 15 avril 2017]. Disponible à l'adresse : <http://www.anaphore.eu/web-des-donnees-et-archives-quel-interet/>
- POUPEAU, Gautier. Quel événement !? Ou comment contextualiser le triplet. *Les petites cases* [en ligne]. 2010. [Consulté le 21 avril 2017]. Disponible à l'adresse : <http://www.lespetitescases.net/quel-evenement-ou-comment-contextualiser-le-triplet>
- POUPEAU, Gautier. Petite histoire du Web sémantique. *Les petites cases* [en ligne]. août 2011. [Consulté le 3 mai 2017]. Disponible à l'adresse : <http://www.lespetitescases.net/petite-histoire-du-web-semantique>
- ROCHKIND, Jonathan. Linked Data Caution. *Bibliographic Wilderness* [en ligne]. 23 novembre 2015. [Consulté le 31 juillet 2017]. Disponible à l'adresse : <https://bibwild.wordpress.com/2015/11/23/linked-data-caution/>
- SAJUS, Bertrand et LEROI, Marie-Véronique. Le développement du web des données culturelles. *I2D – Information, données & documents*. février 2016. Vol. me 53, n° 2, pp. 46-47.
- SIBILLE, Claire. Archives et Données liées : nos outils sont-ils prêts à compléter la Toile ? *Modernisation et archives* [en ligne]. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://siaf.hypotheses.org/344>
- SIBILLE, Claire. Les potentialités du Web sémantique pour le partage du travail des historiens... et des archivistes ! *Modernisation et archives* [en ligne]. [Consulté le 17 avril 2017]. Disponible à l'adresse : <http://siaf.hypotheses.org/170>
- SIBILLE, Claire. Un projet de modélisation pour la description archivistique :

présentation des travaux en cours du Conseil international des archives. *Modernisation et archives* [en ligne]. [Consulté le 19 avril 2017]. Disponible à l'adresse : <https://siaf.hypotheses.org/329>

- VILLEMINOZ, Jérôme. Le web de données à la BNF. *ABES : Arabesques* [en ligne]. N° 83. [Consulté le 2 mai 2017]. Disponible à l'adresse : <http://www.abes.fr/Publications-Evenements/Arabesques/Arabesques-n-83>
- YAKEL, Elizabeth. Archival Representation. *Archival Science*. 2003. pp. 1-25. DOI 10.1007/BF02438926.
- Vade-mecum pour les producteurs de données. *Ministère de la Culture et de la Communication* [en ligne]. [Consulté le 18 août 2017]. Disponible à l'adresse : <http://www.culturecommunication.gouv.fr/Divers/Harmonisation-des-donnees-culturelles/Actualites/Vade-mecum-pour-les-producteurs-de-donnees>



UFR SHA

Mention Information-Communication
Spécialité Documentation

Année universitaire 2016-2017

**Le potentiel des technologies du web sémantique pour
l'archivage dans un contexte de gouvernance de l'information :
*le cas du Système d'Archivage Électronique Mutualisé (SAEM) girondin***

Annexes

Mémoire pour l'obtention du Master esDOC

Présenté par

Alexander Felton

Le 22 septembre 2017

Sous la direction de

Monsieur David Guillemin
Université de Poitiers



Sommaire des annexes

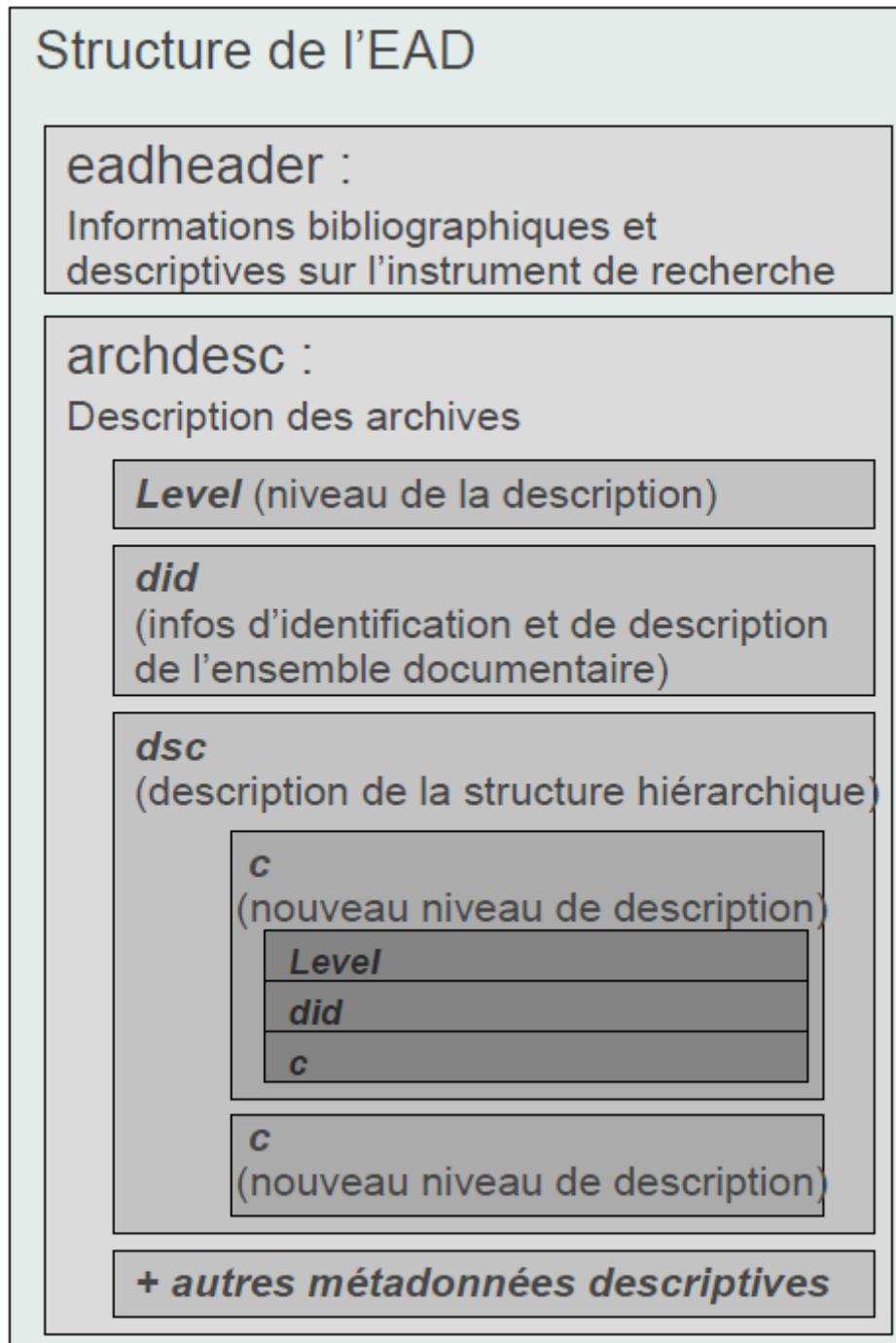
Sommaire des annexes.....	2
Annexe 1 : la théorie des trois âges en archivistique.....	3
Annexe 2 : la structure de la DTD EAD (Encoded Archival Description).....	4
Annexe 3 : le paysage normatif de l’archivage électronique.....	5
Annexe 4 : le modèle OAIS.....	6
Annexe 5 : le « layer-cake » du web sémantique, ou Semantic Web Stack.....	7
Annexe 6a : grille d’entretien (version archivistes).....	8
Annexe 6b : grille d’entretien (version prestataire).....	10
Annexe 7a : transcription de l’entretien avec Pascal Geneste.....	12
Annexe 7b : transcription de l’entretien avec Delphine Jamet.....	23
Annexe 7c : transcription de l’entretien avec Amélie Antoine.....	33
Annexe 7d : transcription de l’entretien avec Damien Dudezert.....	46
Annexe 7e : transcription de l’entretien avec Sylvain Thénault.....	51
Annexe 7f : transcription de l’entretien avec Agnès Vatican.....	57
Annexe 8 : les fonctionnalités du module référentiel du SAEM girondin.....	68
Annexe 9 : la vision modulaire propre au SAEM girondin.....	69
Annexe 10 : le positionnement du module référentiel dans l’architecture du SAEM.....	70
Annexe 11 : typologie des métadonnées du module référentiel.....	71
Annexe 12 : les modules du SAEM girondin et le cycle de vie des archives.....	72
Annexe 13 : articulation des modules du SAEM avec le modèle OAIS.....	73
Annexe 14 : répartition des rôles pour la mise en production du SAEM girondin.....	74
Annexe 15 : l’articulation des différentes couches du modèle de données du SAEM.....	75
Annexe 16 : les user stories issues du backlog de Logilab.....	76
Annexe 17 : le planning du projet SAEM.....	78
Annexe 18 : la vision fonctionnelle cible du SAEM.....	79

Annexe 1 : la théorie des trois âges en archivistique

ÂGE	VALEUR	DURÉE MOYENNE	LOCAUX
« Administratif »	Primaire maximale (usage quotidien)	5 ans	Bureaux
Intermédiaire	I Primaire réduite	5 + 5 = 10 ans	Dépôt intermédiaire proche de l'administration (« petit dépôt »)
	II Primaire faible (référence occasionnelle)	10 + 20 = 30 ans	
	III Secondaire potentielle (attente de l'archivage définitif)	30 + 20 = 50 ans	
« Historique »	Secondaire maximale	Perpétuelle	Dépôt d'archives proprement dit

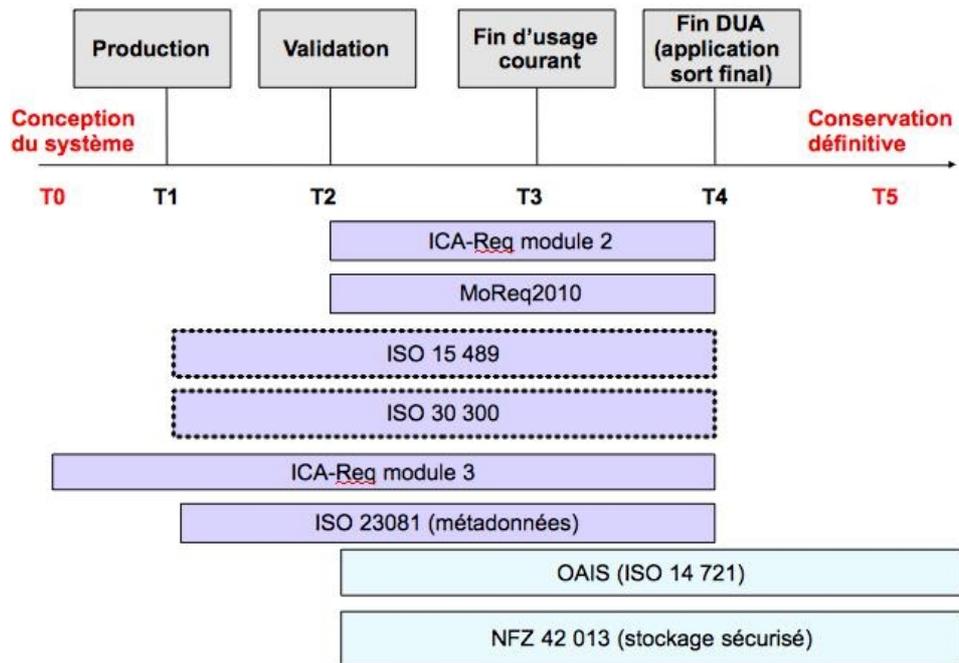
Source : KERN, Gilliane, HOLGADO, Sandra et COTTIN, Michel. Cinquante nuances de cycle de vie. *Les Cahiers du numérique*, 17 juin 2015, Vol. 11, n° 2, pp. 37-76.

Annexe 2 : la structure de la DTD EAD (Encoded Archival Description)



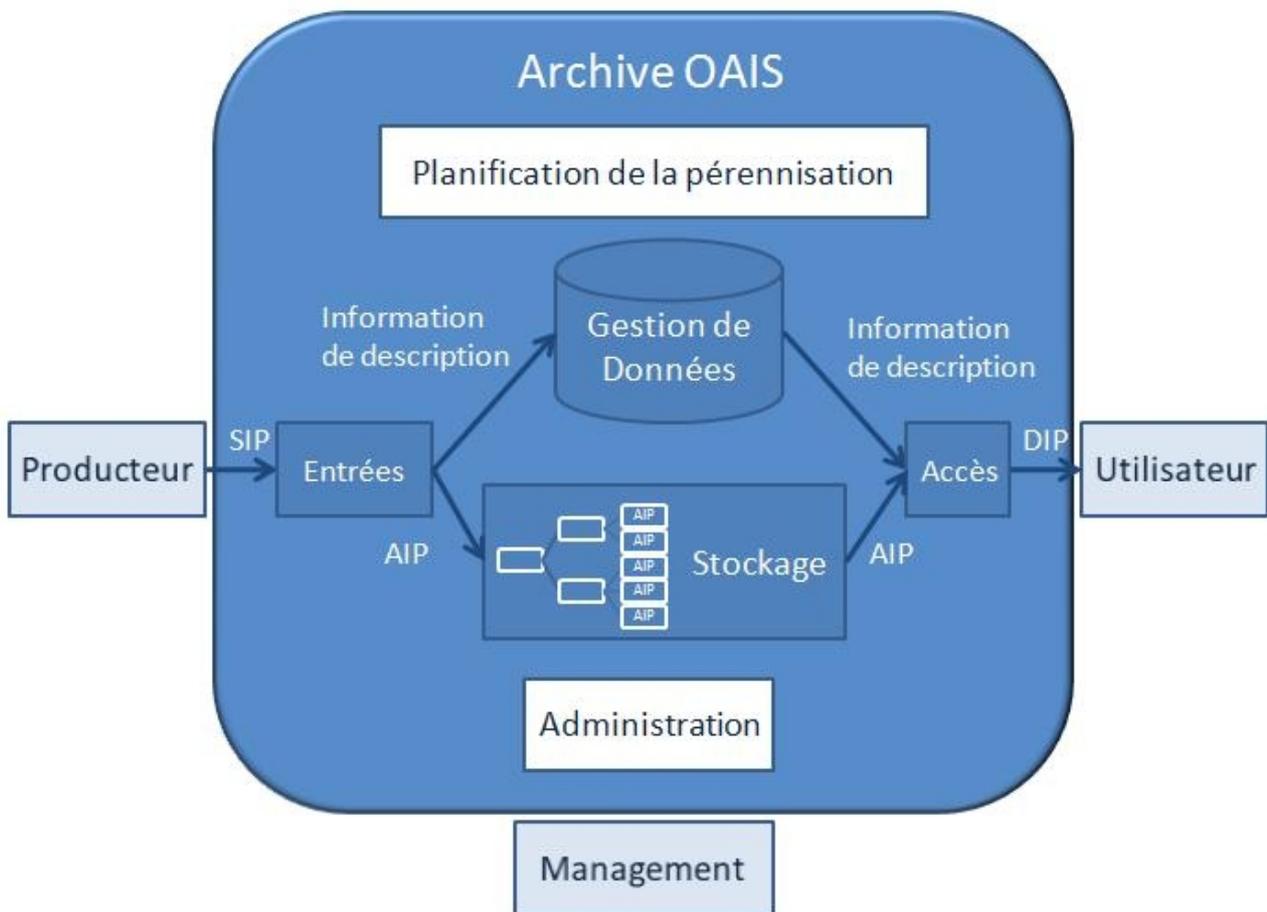
Crédits : Lorène Bécharde. Centre Informatique National de l'Enseignement Supérieur, 2013.

Annexe 3 : le paysage normatif de l'archivage électronique



Source : <http://references.modernisation.gouv.fr/archivage-numerique>

Annexe 4 : le modèle OAIS

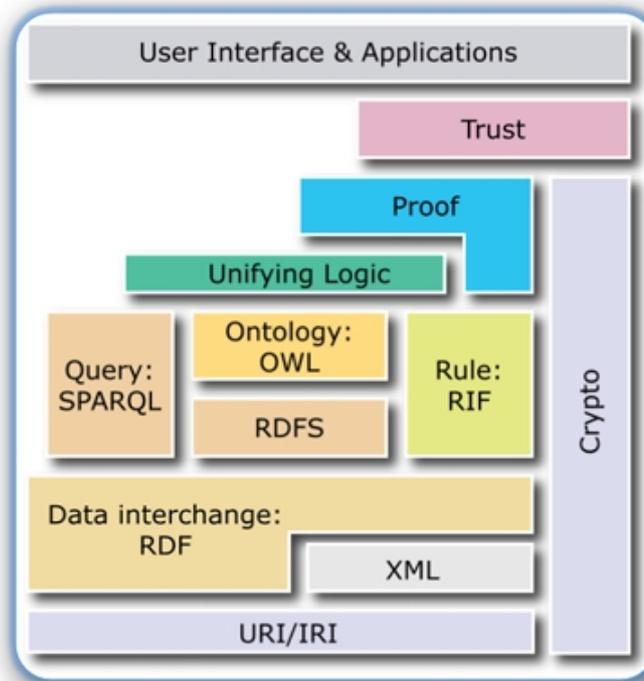


Source : Le modèle de référence OAIS. archivEngines [en ligne]. 7 décembre 2011.

[Consulté le 3 mai 2017]. Disponible à l'adresse :

<https://archivengines.wordpress.com/2011/12/07/le-modele-de-reference-oais/>

Annexe 5 : le « layer-cake » du web sémantique, ou Semantic Web Stack



Source : <https://www.w3.org>

Annexe 6a : grille d'entretien (version archivistes)

- 1) Parcours professionnel et fonctions de l'interviewé :
 - a) Quelles études avez-vous suivies ? Quel est votre parcours ?
 - b) Quel(s) diplôme(s) et/ou concours avez-vous obtenus ?
 - c) Quel est votre positionnement au sein de la structure où vous travaillez ? Quelles sont vos interactions avec les autres services ? Quelles sont vos principales fonctions et tâches au quotidien ?

- 2) Rôle dans la mise en œuvre du projet :
 - a) Depuis quand participez-vous au projet SAEM ?
 - b) Comment avez-vous été amené à travailler sur ce projet ?
 - c) Avec qui êtes-vous amené à collaborer et à échanger principalement ? Comment se déroulent ces échanges ? Les différentes cultures professionnelles et visions compliquent-elles les échanges ?
 - d) Comment percevez-vous l'équilibre des forces en présence dans les différentes instances du projet ? Cet équilibre a-t-il évolué depuis le lancement du projet ?

- 3) Archivage électronique et bouleversement des pratiques professionnelles :
 - a) Avez-vous ressenti une modification des pratiques ces dernières années, avec l'avènement de l'archivage électronique et avec l'effort de numérisation des archives papier et de leur mise en ligne sur Internet ? Dans quelle mesure ?
 - b) Que pensez-vous de la traduction de certains grands principes de l'archivage lors de leur application à l'environnement numérique ? Considérez-vous que les nouveaux standards qui s'imposent (norme 30300, modèle OAIS, etc.) correspondent aux pratiques professionnelles actuelles ?
 - c) Les instruments de recherche, dans leur forme actuelle, constituent-ils toujours selon vous un outil pertinent ? Sont-ils toujours adaptés aux usages des différentes catégories de publics ?
 - d) Pensez-vous qu'on puisse aujourd'hui parler d'une convergence des approches (records management, archivistique « traditionnelle », informatique) en matière d'archivage ?
 - e) Avez-vous suivi des formations dédiées à ces aspects ?

- 4) Estimation du degré de familiarité avec les concepts liés au web sémantique :
 - a) Que connaissez-vous des différentes technologies du web sémantique ? Comment avez-vous été amené à vous intéresser ou à vous confronter à ces questions (Formation ? Lectures ? Projet professionnel ?...)
 - b) Comment distinguez-vous cette notion web sémantique du concept de web des données ?
 - c) Comment envisagez-vous la place du module référentiel dans le projet SAEM ? Quels sont ses objectifs et comment percevez-vous ses fonctionnalités ?
 - d) De quelle manière concevez-vous la compatibilité entre ces technologies et la tradition de description et de classement en archivistique ? Quelle influence aura selon vous le nouveau modèle conceptuel de description archivistique ?
 - e) Questionner la perception de leur efficacité et de leur potentiel : par quel biais ces

outils peuvent-ils servir aux professionnels et aux publics ? Considérez-vous que les professionnels et les publics sont prêts à accueillir ces changements ? Certaines catégories de publics vont-elles bénéficier davantage de ces évolutions ?

- 5) Perception des différents enjeux du web des données et de son intérêt pour les données issues du domaine culturel et des archives :
 - a) Avez-vous connaissance de différents projets ou ressources utilisées dans le domaine de la culture qui reposent sur le web des données ?
 - b) Quel a été l'impact de la feuille de route du Ministère de la culture concernant les métadonnées culturelles et la transition web 3.0 ? Dans les faits, avez-vous ressenti de nouvelles impulsions ?
 - c) Comment percevez-vous les fonctionnalités du site web d'un service d'archives et les utilisations qui en découlent à l'heure actuelle ?
 - d) Quels sont les enjeux de l'exposition en ligne des données issues du traitement de documents d'archives numériques pour les chercheurs aujourd'hui ? Comment percevez-vous le potentiel de ce virage de la donnée qui s'amorce et notamment du web des données ? Comment cela s'articule-t-il avec le mouvement des humanités numériques ?

Annexe 6b : grille d'entretien (version prestataire)

- 1) Parcours professionnel et fonctions de l'interviewé :
 - a) Quelles études avez-vous suivies ? Quel est votre parcours ?
 - b) Quel(s) diplôme(s) et concours avez-vous obtenus ?
 - c) Quel est votre positionnement au sein de la structure où vous travaillez ? Quelles sont vos principales fonctions et tâches au quotidien ?

- 2) Interactions avec les archivistes et autres professionnels de l'information
 - a) Avec qui êtes-vous amené à collaborer et à échanger principalement ? A combien de projets similaires au SAEM avez-vous participé ? Comment se déroulent ces échanges ?
 - b) Les différentes cultures professionnelles et visions compliquent-elles ces échanges ?
 - c) Échangez-vous davantage avec les informaticiens impliqués dans le projet ou bien avec les archivistes ?

- 3) Appropriation du vocabulaire des archivistes et des problématiques de l'archivage électronique
 - a) Quels sont les éléments qui ont facilité l'appropriation de ces concepts par l'ensemble du COPROJ ?
 - b) Avez-vous pu observer des rapprochements entre les procédures et la conception globale de l'archivage électronique entre les archivistes et informaticiens, à travers le prisme de la mise en place du projet SAEM ?
 - c) Avez-vous suivi des formations dédiées à l'archivage électronique, aux métadonnées issues du traitement de documents d'archives ?

- 4) Intégration des technologies du web sémantique dans les outils et mise en place de modèles de données dédiés à l'archivage :
 - a) Comment avez-vous été amené à vous intéresser ou à vous confronter à ces technologies (Formation ? Lectures ? Projet pro ?...)
 - b) Quel est l'intérêt de ces technologies pour les différents produits que vous éditez ? Qu'est-ce qu'elles permettent et comment percevez-vous leur valeur ajoutée ?
 - c) D'après vous, ce potentiel est-il perçu par les archivistes et vous ont-ils fait part de leur volonté de développer ces aspects ? Quelles sont leurs attentes par rapport au potentiel de CubicWeb ?
 - d) Comment concevez-vous la place du module référentiel dans le projet SAEM et plus particulièrement l'impact du recours aux technologies du web sémantique ?
 - e) De quelle manière ces outils peuvent-ils servir aux professionnels et aux publics ? Considérez-vous que les professionnels et les publics sont prêts à accueillir ces changements ?

- 5) Perception des différents enjeux du web des données et de son intérêt pour les données issues du domaine culturel et des archives :
 - a) Dans quelle mesure êtes-vous familier des différents projets ou ressources utilisées dans le domaine de la culture qui reposent sur le web des données ?
 - b) Quel est l'intérêt, d'après vous, de la mise en relation de vocabulaires d'horizons divers dans le cadre de l'indexation des documents d'archives ?

- c) Comment s'est organisée la conception du modèle des données propre au projet SAEM ? Quels sont les principaux enjeux qui sous-tendent ses différentes facettes ?

Annexe 7a : transcription de l'entretien avec Pascal Geneste

Entretien avec Pascal GENESTE, responsable du service collecte et traitement des Archives départementales de la Gironde et membre des COTECH et COPIL du projet SAEM.

Réalisé dans son bureau, en début de matinée, le 14 juin 2017.

Alors, déjà, est-ce que vous pouvez commencer par me parler un peu de vous et de votre parcours ?

Oui, alors, je m'appelle Pascal GENESTE. Mon parcours, il est assez classique en fait, au poste que j'occupe aujourd'hui, puisque j'ai commencé, après le bac, à faire une prépa pour préparer l'école des chartes. J'ai grandi dans le Lot-et-Garonne et je suis parti à Paris faire ma prépa, à l'école des chartes, donc la scolarité dure 4 ans, j'ai préparé une thèse en Histoire moderne, thèse d'Histoire religieuse d'ailleurs : la biographie d'un évêque au 18^e siècle. Et puis, ensuite, une fois mon diplôme d'archiviste-paléographe en poche, j'ai présenté le concours d'entrée à l'école nationale du patrimoine, qui s'appelle maintenant l'institut national du patrimoine, mais qui était depuis 5-6 ans l'école, en quelque sorte, professionnalisante, que devaient passer... par laquelle les conservateurs, les archivistes en tout cas, devaient passer pour obtenir un poste de conservateur. Donc, j'ai fait partie des premières promotions de l'école du patrimoine, la scolarité durait 18 mois, et c'est une scolarité faite à la fois de stages théoriques et pratiques, cours théoriques plutôt et stages pratiques, à l'issue desquels j'ai été nommé en premier poste comme responsable des Archives de la Marine, au château de Vincennes, dans un service qui ne s'appelait pas encore le service historique de la Défense, puisqu'il n'était pas mutualisé comme il l'a été quelques années plus tard, mais c'était le service historique de la Marine. Je suis resté quatre ans et ensuite, donc, j'ai fait le choix d'aller aux Archives nationales où l'on m'a proposé un poste de responsable des archives des présidents de la Ve République, qui globalement, en fait... je travaillais sur les archives des hommes politiques de la fin du 20^e siècle. J'y suis resté une dizaine d'années et je suis arrivé à Bordeaux en 2011, juste après l'installation dans ce nouveau bâtiment, dans lequel nous sommes, du cours Balguerrie-Stuttenberg, quelques mois après, et donc ça fait presque six ans que je suis aux fonctions de directeur adjoint responsable de la collecte et du traitement des fonds en Gironde.

Et justement, ce poste-là, qu'est-ce que ça implique en termes de fonctions, d'activités au quotidien ?

Alors, par rapport à mes fonctions précédentes aux Archives nationales, où on était plus spécifiquement responsable scientifique, en quelque sorte de fonds que l'on gérait en propre, j'allais dire, de la collecte à la valorisation, sans forcément d'appui technique, ici c'est beaucoup plus un poste administratif, de responsabilité, puisqu'on est dans un service d'archives départemental qui est considéré comme l'un des plus importants en termes de kilométrage conservé et en termes d'effectif. Donc ça implique évidemment des responsabilités en termes de gestion de ressources budgétaires, financières, de ressources humaines, que les postes précédemment occupés. Des responsabilités qui ont été décuplées, j'allais dire, assez rapidement, puisqu'au bout d'un an, le directeur a été nommé RAC ? en Pays de la Loire, et donc on m'a demandé d'assurer les responsabilités de directeur par intérim, de ces archives départementales, et donc là pendant quelques mois j'ai dû assumer à la fois les fonctions de directeur mais aussi de responsable de la collecte et du traitement, sinon d'assurer les fonctions des deux côtés.

Et, ça veut dire que vous interagissez principalement avec quelles... avec qui, dans les différents projets ?

Bon, le service, vous savez comment il est composé, il est composé principalement, et ça a été une des missions qui m'a été demandée de mener à bien au moment de mon intérim, de restructurer

deux services cœur de métier qui sont : service des publics, puisqu'on est avant tout un service public, au service des administrations publiques mais aussi au service des usagers. Donc ce service des publics, il est essentiel dans notre dispositif, et il est composé aujourd'hui, vous le savez, de deux bureaux : un bureau d'accueil, en quelque sorte, du public, gestion de la salle de lecture, gestion du service éducatif et puis un bureau plus spécifiquement chargé des recherches et de la valorisation des fonds : recherches sous toutes ses formes, ça peut être des recherches pour l'administration, ça peut être des recherches pour les généalogistes, comme ça peut être des recherches pour les chercheurs, les historiens, amateurs ou professionnels. On ne fait aucune distinction, d'ailleurs, entre les usagers, on ne doit pas en faire, il y a un principe d'équité, d'égalité de traitement des usagers. Donc ça c'est l'un des services cœur de métier de notre dispositif, et le second, celui dont je m'occupe en propre, c'est celui qui est chargé à la fois de la prospective et de la collecte des fonds, qui est chez nous extrêmement importante puisque c'est de l'ordre d'un kilomètre à deux kilomètres par an, et depuis au moins une quinzaine d'années, c'est pas quelque chose d'exceptionnel, mais ensuite, une fois que les archives sont collectées, on est chargés également de leur traitement intellectuel et matériel, par le biais d'un bureau, qui physiquement se trouve avoir été installé dans notre site historique de la rue d'Aviau : bureau du classement et de la conservation. La terminologie, qui est employée depuis maintenant quatre ans, indique bien que nous avons voulu recentrer ces deux services cœur de métier sur les principales missions, les principales fonctions des archives, qu'on appelle les 4C : collecte, classement, conservation, communication, auxquelles on peut rajouter la valorisation, mais on a vraiment voulu recentrer l'organisation structurelle de ces deux services cœur de métier sur ces missions principales alors qu'auparavant, quand je suis arrivé encore, il y a 6 ans, donc ça ne fait pas si ancien 6 ans, on avait une structuration d'ailleurs bien plus grande, là où on a quatre bureaux aujourd'hui, on en avait sept, quand je suis arrivé il y a 6 ans, avec une structure qui ne se répondait pas forcément l'une par rapport à l'autre, puisque les quatre bureaux du service collecte et traitement étaient organisés thématiquement. Il y avait un bureau qui s'occupait des archives du Conseil général, un autre bureau s'occupait des archives de l'État, un autre qui s'occupait des archives notariales et privées et un quatrième, géré par Delphine, plutôt des archives contemporaines. Donc, évidemment, chacun de ces quatre bureaux, qui étaient d'ailleurs sous-dimensionnés, par rapport aux missions qui étaient les leurs, s'occupait aussi bien de collecte que de traitement, sans qu'il n'y ait forcément une cohérence d'ensemble à tout cela. Donc, c'est l'une des principales missions d'un responsable, comme nous le sommes avec George Cuer, que d'organiser ces missions, afin qu'elles se répondent les unes les autres, mais aussi qu'elles soient menées le plus efficacement possible, et avec le moins de dispersion possible.

D'accord. J'aurais voulu revenir sur votre rôle dans le cadre du projet SAEM, étant donné qu'on a eu l'occasion de se croiser lors d'une réunion du COTECH, le comité technique. Comment est-ce que vous avez été amené à travailler sur ce projet ?

Alors, le projet de système d'archivage électronique mutualisé a été l'un des projets phares qui était présenté en préfiguration au moment où j'ai candidaté sur ce poste-là. La Gironde, de l'extérieur, je ne connaissais pas plus que ça le service d'archives départementales de la Gironde, si ce n'est par la réputation qu'il pouvait avoir acquise au fil du temps. C'était un service qui était plutôt axé, et un peu exemplaire en même temps, sur la collecte. Tout d'abord parce que la conservatrice qui m'a précédée est restée presque un quart de siècle et venait de ce que l'on appelle les archives auprès des ministères, à Paris, qui sont des missions qui, vous le savez, sont là en amont du rôle des archives nationales pour justement convaincre les administrations centrales, à la fois les convaincre de verser leur production documentaire mais aussi organiser la mise en œuvre de cette collecte-là. Donc c'était une personne qui avait une expérience d'administration centrale assez rodée, assez pionnière, et qu'elle a voulu essayer développer, à juste titre, à un niveau plus territorial. Donc je venais sur un terrain où la collecte était considérée comme une entité très importante, ce qui n'est

pas forcément le cas dans tous les services d'archives puisque certains services souffrent de leurs locaux, de manière très pragmatique, et sont dans des bâtiments souvent assez anciens et saturés. Donc, quand vous êtes dans un bâtiment assez ancien et saturé, vous pouvez mettre en œuvre des actions de valorisation des fonds, des actions de conservation préventive, poursuivre les tâches de classement, mais vous êtes quand même un peu bloqué dans les missions de collecte. Ce n'était pas du tout le cas en Gironde, puisque comme je le disais tout à l'heure en préambule, j'arrivais à la fois dans un service où la collecte était depuis à peu près un quart de siècle une mission phare et d'autre part dans un service qui venait de se doter d'un nouveau bâtiment qui lui permettait d'envisager une collecte future importante. Et en même temps à une période où, depuis une dizaine d'années, on assiste à une forme de transformation de cette collecte-là, avec une production hybride de dossiers papier, qu'on voit continuer à fleurir dans les administrations, même celles qui prônent le zéro papier, et cette production électronique, dont on a bien conscience qu'elle est importante et qu'il faut la prendre en compte. Donc, quand j'ai postulé, l'un des premiers projets qu'il m'a été demandé de suivre, c'est ce projet de mise en place d'une plateforme d'archivage électronique, mutualisé à l'époque avec la région Aquitaine, la ville de Bordeaux et ce que l'on n'appelait pas encore la métropole mais la communauté urbaine. J'ai été tout de suite intégré au comité de pilotage et assez rapidement, avec la chargée de projet à l'époque qui était Véronique Pisani, pour essayer de travailler sur la première phase d'étude. Ensuite, le groupe projet s'est un peu transformé, puisque le système est monté en puissance et il a fallu également associer rapidement Delphine à Véronique et progressivement, c'est Delphine qui a pris le pas, et qui a intégré pleinement et totalement le COPROJ, le comité projet, tout en me rendant régulièrement compte de l'avancée des réflexions qui sont menés, et mon rôle consiste plus à participer au comité de pilotage et au comité technique lorsqu'ils se réunissent, et d'essayer de promouvoir à l'extérieur le projet, notamment en essayant de lever des subventions auprès du SIAF chaque année, en participant à l'appel d'offre AD-ESSOR, et de valoriser le projet sur le plan de la communication, de l'information, comme on va le faire par exemple demain, auprès de collègues de la Corrèze ou qu'on le fera début juillet auprès de l'ensemble des collègues départementaux de la nouvelle région Nouvelle Aquitaine.

Donc, sur ce projet-là, vous avez davantage un rôle administratif et décisionnel, finalement.

Plutôt, voilà.

Donc, vous n'êtes pas...

Et comment c'est structuré finalement au sein du service, avec un binôme, en quelque sorte, avec Delphine et moi, puisque c'est mon adjointe directe et celle à qui je confie les dossiers transversaux que je n'ai pas le temps de suivre pleinement mais aussi le partage d'information au quotidien. C'est son appétence pour la question et son expérience, depuis une quinzaine d'années, de développement de systèmes d'information, on va le dire rapidement comme ça, me permet de m'appuyer sur quelqu'un d'extrêmement solide et qui comprend bien la problématique.

Finalement, vous n'êtes pas directement en lien avec les équipes du côté DSI, les informaticiens ?

Alors, on peut l'être mais souvent on est en lien avec Hélène Lapouble ou Phillippe Gazengel, qui lui aussi au départ était intégré au COPROJ et qui maintenant est plutôt sur le côté organisationnel ou administratif des choses. Donc on est en lien par ce biais-là ou par le biais de points qu'on peut faire la directrice des systèmes d'information. Là, on travaille énormément, pour ne pas vous le cacher, on travaille surtout, dans le cadre de ce projet, sur la mise en place de la convention, donc le dialogue avec le partenaire historique qui reste encore en jeu, c'est-à-dire la ville et la métropole, puisque ces collectivités ont mutualisé leur fonction support. Donc on travaille essentiellement à l'élaboration et à la conception de la convention pour la mise en place de la phase 3 et comme vous le savez, ce travail n'est pas sans écueil et sans lourdeur administrative.

J'aurais voulu également aborder le rapport entre, justement, cet avènement de l'archivage

électronique et les pratiques traditionnelles de l'archivistique.

Alors moi je suis d'une génération qui n'a pas, comme vous, grandi avec l'informatique, et je le vois aussi à travers mes enfants. Donc je suis d'une génération qui pense, de toute manière, qu'il faudra au moins une bonne trentaine d'années avant que cette mutation entre la production traditionnelle, la production papier, qui est une tradition très ancienne, ne soit pleinement, peut-être pas pas aboutie d'ailleurs, mais plutôt pleinement prise en compte. Il y a une dizaine d'années, on pensait que l'on allait pouvoir passer, comme ça, d'une archivistique que l'on pourra qualifier d'antan, mais qui reste actuelle, à une archivistique ne gérant essentiellement plus l'électronique, les métadonnées. Bon, on s'aperçoit quand même, 10-15 ans après, qu'on reste très largement sur cette archivistique traditionnelle, et que ceux qui étaient chargés de la collecte de certaines administrations publiques continuent à l'être sans pour autant être impliqués dans de la collecte de documents électroniques. C'est d'ailleurs l'une des problématiques que l'on aura à résoudre dans les années à venir. On génère dans les administrations beaucoup de vrac électronique, beaucoup de ressources qu'il faudra, si on le peut, traiter un jour, mais pour l'instant qu'on laisse plutôt de côté, parce qu'on a pas les moyens de les prendre en charge, sinon des moyens très basiques de sauvegarde, sur disques durs ou serveurs, de certains ensemble. Ça a été le cas, notamment pour ce qu'on a considéré comme très important, je pense à la messagerie électronique d'un directeur de cabinet du préfet, je pense à certains ensembles de ce type, qui sont finalement une goutte d'eau par rapport à la production extrêmement massive de l'administration. L'autre élément, c'est l'articulation entre ce qui se passe au niveau central, puisque finalement nous gérons des services déconcentrés de l'État, mais qui ont une tutelle au niveau central très importante et encore très prégnante dans un pays centralisateur jacobin comme l'est la France, c'est une évidence. Donc, il y a à prendre en compte cette articulation entre ce qui se passe au niveau central et ce qui peut être développé au niveau territorial et on le voit d'ailleurs dans les derniers échanges qu'on a pu avoir dans le comité de pilotage du SAEM, où on commence à se dire qu'il faudrait peut-être rencontrer nos interlocuteurs qui suivent, à Paris, le projet VITAM, qui est un projet interministériel de prise en compte de l'archivage électronique, pour voir si l'on va pouvoir articuler nos propres réflexions et l'aboutissement de ces réflexions, c'est-à-dire la mise en œuvre et l'expérimentation de notre système, avec ce qui se passe au niveau central. Donc, je ne dis pas que chacun se renvoie à la balle finalement, mais on voit que ce ne sont pas des systèmes très simples à modéliser, d'abord, mais ensuite à mettre en œuvre. J'ai parlé un petit peu de mon parcours, j'étais il y a une vingtaine d'années au ministère de la Défense, à la fin des années 90, début des années 2000, on parlait déjà d'archivage électronique et de développer des systèmes d'archivage électronique, qui n'ont jamais abouti, et sur lesquels le ministère de la Défense, via VITAM, avec maintenant le ministère de la Culture et le ministère des affaires étrangères, essayent de faire aboutir.

Et d'après vous, d'où viennent ces difficultés à faire aboutir ces projets ?

Eh bien parce que le sujet est complexe. Si depuis 2012 en Gironde, on travaille déjà sur élaboration d'un schéma et la mise en œuvre de ce schéma sur simplement quatre flux, c'est pas... Quatre flux qui sont peut-être représentatifs de ce que ces quatre entités historiques, qui ont voulu se réunir sur ce sujet, en les considérant un peu essentielles. Mais ces quatre flux ne représentent évidemment pas l'ensemble de la production électronique que peut produire l'administration, puisque l'un des grands absents de ces grands partenaires, c'est quand même l'État. Or, l'État, pour nous, en proportion et en statistique, ça représente quand même 90 % de notre collecte. Donc finalement, on ne travaille que sur quatre flux, des 10 % de la collecte traditionnelle. Donc, je ne sais pas ce que ces quatre flux représentent en pourcentage sur ces 10 %, mais on ne travaille aujourd'hui que sur une toute petite partie, finalement, de la production électronique de l'administration française, même dans un territoire comme le nôtre et évidemment, il va bien falloir élargir une fois que les premiers systèmes seront en place, élargir ce champ, c'est pour ça que je parlais d'une trentaine ou quarantaine d'années, pour progressivement arriver à absorber la totalité de la production

électronique.

Donc, finalement, la difficulté principale, ça serait d'arriver à mettre en place des modèles communs ?

C'est ça, à la fois à mon avis réduire les typologies, qui ont explosé. C'est vrai pour le papier et peut être encore plus pour l'électronique, puisqu'on est là avec un outil qui nous permet de multiplier la production, par le biais simple du copié-collé, mais aussi parce qu'à chaque fois qu'on produit une version supplémentaire d'un document, on n'écrase pas la précédente, on l'ajoute en quelque sorte à la précédente. Du coup, comme l'on avait assisté, dans la production papier, à une vraie explosion documentaire depuis la fin de la seconde guerre mondiale, puisqu'on considère souvent, dans nos dépôts d'archives départementales que 80 % de ce que l'on conserve est postérieur à la seconde guerre mondiale. Comme on a assisté à un élargissement considérable de la notion de patrimoine, et aujourd'hui tout est patrimoine, et à l'intérieur de cette notion-là, il y a les archives évidemment mais aujourd'hui tout est archive. On risque d'assister, il faut être un peu prospectif, mais on risque d'assister dans les années à venir à une réduction de cette notion-là, à une réduction aussi certainement des typologies documentaires dont je parlais tout à l'heure, et à une réduction du champ d'action des services d'archives, qui ne pourront pas, et des services informatiques, qui sont maintenant pleinement associés à cette réflexion, qui ne pourront pas absorber la totalité de ce qui est produit, et qui seront obligés de mettre, en quelque sorte, des gardes-fous. On parle souvent de l'utilisation pleine et entière des boîtes fonctionnelles plutôt que des boîtes personnelles, etc. Bon, ça c'est un accompagnement, aussi, de l'utilisateur, et des services administratifs, pour arriver à contrôler tout ce qui a été mis en place ces dernières années, mais sans forcément penser à ce que cela deviendrait. J'en veux pour preuve les différents systèmes de gestion électronique des documents qui ont pu se développer pêle-mêle, sans jamais prendre en compte la question de l'archivage des données qui étaient comprises dans ces systèmes de gestion là. Donc, on est passé progressivement d'un développement massif de systèmes de gestion pour répondre à la demande de tel ou tel service, à une deuxième étape qui a constitué à dire : « bon, on ne peut pas continuer à développer des GED pêle-mêle en fonction de tel ou tel besoin, il faut arriver à produire un modèle de GED transverse » et ensuite une troisième étape qui consiste à dire : « ce système de GED transverse, c'est bien joli, mais qu'est-ce qu'on va faire des données qui seront comprises dans ce système de gestion, à la fin de ce que l'on appelle de la durée d'utilité administrative, qui est vraie pour le papier mais aussi pour l'électronique. Une fois qu'on a plus un usage administratif de ces données-là, qu'est-ce qu'on en fait ?

Est-ce que, du coup, vous pensez qu'il y a une sorte de convergence entre les différentes approches de l'archivage entre, justement, tout ce qui est GED et records management et l'archivage traditionnel et la vision assez spécifique qu'ont les informaticiens ?

Alors, il faudrait qu'il y en ait de plus en plus. C'est pour cela que nous avons pour ambition, dans les mois et années à venir, de faire monter en compétence les nombreux agents que nous avons, et notamment ceux qui s'occupent de la collecte et du traitement, sur ces problématiques-là. On l'a dit tout à l'heure, aujourd'hui à part Véronique Pisani et Delphine Jamet, on n'est pas nombreux à s'être interrogé et avoir plongé dans les réflexions visant à faire aboutir ce SAE. Donc l'idée, c'est qu'il y ait une appropriation de beaucoup plus d'agents sur cette question-là, et sur le système qui aura été développé, et sur son utilisation, pour que progressivement cela infuse auprès des producteurs de documents, qui eux-mêmes peuvent dans leurs propres services, organismes et administrations avoir des sollicitations. Il y a des organismes beaucoup plus en pointe que d'autres, et donc il faut être en capacité à la fois de répondre à ces services en pointe mais aussi à ceux qui ne le sont pas du tout et qui découvrent ces problématiques-là. Donc, il y a une corrélation évidente entre la production papier et la production électronique, qui se complètent. Donc par ce biais-là, il faut qu'il y ait une forme d'action partagée entre le traitement du papier et des données

électroniques. Donc, la montée en compétence des agents, c'est une chose, ensuite la priorisation de certaines actions, pour le traitement de ces ensembles hybrides, ça en est une autre. Vous le savez, dans l'archivistique papier, on a beaucoup mis en place des tableaux de gestion, qui sont là pour décrire précisément la production documentaire d'un organisme ou d'une administration, par typologies, auxquelles on met en regard des durées d'utilité administratives et des sorts finaux. Il faut évidemment que l'on intègre à cette organisation en tableaux de gestion, toutes les données électroniques qui aujourd'hui peuvent avoir été créées ou qui peuvent être créées dans les moi ou années à venir.

Et est-ce que vous, vous avez suivi des formations, justement, relatives à tous ces aspects-là ?

Alors on a des formations, oui, on a pu suivre quelques formations. Souvent, ce sont des formations sur les formats de données. Alors, il y a une dizaine d'années, c'était le XML-EAD, puisque l'on commençait, dans le cadre de notre mission de classement, essentiellement de traitement intellectuel des fonds... Il y a eu une mutation, aussi, entre la production des inventaires papier, traditionnels, que l'on peut continuer à faire dans certains cas, et la constitution de bases de données, d'instruments de recherche électroniques, le développement aussi des sites internet, qui devaient intégrer cette production d'instruments de recherche électroniques... qui ne date que d'une dizaine d'années, pas plus, donc dès qu'on a eu à aborder ces problématiques, il y a eu des formations qui nous ont permis de nous familiariser un petit peu. Les formations, c'est comme tout, il faut ensuite avoir l'occasion de les mettre en pratique, de les mettre en œuvre. Donc, essentiellement une formation sur les formats, sur les formats d'échange, sur les progiciels de gestion qui pouvaient se développer, etc.

Vous avez mentionné les instruments de recherche, justement, est-ce que d'après vous, ils constituent toujours un outil pertinent, finalement, à la fois pour les archivistes qui les construisent et pour les publics ?

Bah, pour l'instant, on considère que sans instrument de recherche, une archive, ou une donnée, ne peut pas être retrouvée. Donc, on peut appeler ça comme on veut, pour moi un instrument de recherche, je parlais tout à l'heure de l'élargissement de certaines notions comme la notion de patrimoine et d'archives, la notion d'instrument de recherche peut être très large. Un instrument de recherche, ce n'est pas forcément un inventaire analytique qui va décrire, à la pièce, un document ou une donnée, ça peut être une description, au niveau haut, d'un ensemble de données, qui peuvent être très larges. D'ailleurs, depuis trois ou quatre ans, dans mon service, on a essayé de développer, avec Delphine, justement pour décrire les versements contemporains, dont je disais tout à l'heure qu'ils représentent la majeure partie des fonds que nous conservons, donc pour les décrire dans leur globalité, alors qu'ils représentent peut-être 60 ou 65 km linéaires, il faut des instruments de recherche, des états généraux de versements contemporains thématiques, donc par producteur. Des instruments de recherche qui les décrivent de manière globale, donc on est parti d'abord de la description, via des notices normalisées, des producteurs en prenant en compte leur histoire, qui a pu évoluer dans le temps, etc., pour ensuite aller décrire tous les versements qu'ils ont pu effectuer depuis leur existence récente. Donc ça c'est un outil de recherche, finalement, un instrument de recherche, qui me paraît effectivement très pertinent, pour permettre à l'utilisateur public, d'avoir des sources contemporaines auxquelles il n'aurait pas accès sans cet outil général. Parce que souvent, et évidemment vous le savez, les versements se font au coup par coup, pas forcément de manière programmée très en amont, et donc il y a une dispersion de la production archivistique qui peut nous être versée, qui n'est pas lisible si on ne la met pas en perspective de manière large. Donc je crois, pour répondre à votre question, que l'instrument de recherche est toujours un outil très pertinent, si on le conçoit essentiellement comme un outil d'aide à la recherche. Alors, évidemment, ces instruments de recherche évoluent dans le temps et on est passé d'une pratique érudite, dans les outils de recherche comme dans le traitement archivistique, à, aujourd'hui, une gestion des flux ou

des masses et une gestion plus globalisée des outils de recherche et plus diversifiée aussi. Pour moi, les cartes mentales qu'on essaie de développer, que Delphine a pu vous montrer, sur certaines typologies documentaires, ou même ensembles de fonds sur une thématique donnée, constituent un instrument de recherche, un outil de recherche. Et il y a deux ans encore, on n'en parlait pas encore, dans le domaine qui est le nôtre, dans le domaine archivistique. Donc il y a forcément à inventer de nouveaux outils de recherche pour accéder... le but c'est quoi ? Le but c'est bien que l'utilisateur puisse en connaître et puisse accéder à l'ensemble des documents que nous conservons, dans la mesure où, effectivement, cet accès autorisé, puisqu'il y a aussi la problématique de la communication des archives publiques, qu'il ne faut pas négliger quand on développe des outils de recherche.

Et justement, au niveau des pratiques de description archivistique, à plusieurs niveaux et avec la hiérarchie que ça impose, est-ce que vous pensez qu'elles sont toujours compatibles avec cet environnement numérique, où l'on accorde de plus en plus d'importance à la donnée ?

Alors, cet environnement numérique, me semble-t-il, a permis une structuration de la description archivistique que l'on n'avait pas forcément auparavant, que l'on n'appliquait pas de manière aussi systématique auparavant. D'abord, les normes internationales de description sont très récentes, elles ont une vingtaine d'années, et elles sont revues régulièrement, puisque là on travaille à une norme générale, Records in Context, qui englobe finalement les quatre normes internationales qui coexistaient auparavant, et dont on usait assez peu dans leur globalité. Il y a essentiellement la norme ISAD(G) pour la description des fonds et la norme ISAAR pour les producteurs, qui étaient peu ou prou adoptées et mises en œuvre, mais il me semble que l'évolution technologique à laquelle on assiste nous contraint, et contraint les archivistes, à utiliser à la fois beaucoup plus les normes et la structuration de la description qu'on ne le faisait auparavant. Il y a certains services qui continuaient des outils de recherche qui n'étaient pas forcément normalisés, pas forcément structurés, aujourd'hui ce n'est plus guère possible si l'on veut apparaître dans un autre cadre que le strict cadre physique dans lequel on travaille. Alors, il y a ça et, évidemment, les opérations de coordination qui sont développées au niveau central et qui ne l'étaient peut-être pas, ou moins, auparavant, et qui permettent de s'inscrire dans cette normalisation et structuration des données dont je parlais. On travaillait hier sur l'indexation des registres matricules, une typologie qu'on va retrouver dans la centaine de services d'archives départementales de France. Si on veut normaliser cette indexation et qu'elle apparaisse dans un portail développé au niveau national, qui s'appelle le Grand Mémorial, et dans lequel vont se retrouver d'autres typologies que celle que j'ai évoqué, on est obligé de normaliser et de structurer sa description et ses données, sinon vous n'entrez pas dans le... vous ne passez pas le portail en quelque sorte, vous restez à la porte. Et donc, c'est une évolution, qu'il faut prendre en compte, dans les pratiques et dans les méthodes de travail des agents, parce que derrière tout ça, il y a de la conception, les concepts qu'on peut développer etc, mais derrière tout ça il y a quand même les agents, les fonctionnaires qui sont là, soit depuis peu soit depuis longtemps, et dont il faut prendre la tâche. Il y a l'accompagnement au changement, qui est une donnée importante et dont on parle, comme ça, en termes très théoriques mais, sur le plan pratique, n'est pas toujours mis en œuvre. Ça, il me semble que c'est quelque chose d'important, en tout cas pour moi qui ai quelques responsabilités en la matière. Mettre des agents sur la conception d'un référentiel, mettre des agents sur la normalisation et la structuration des données etc, c'est important, mais il faut leur expliquer pourquoi et dans quel contexte on le fait, ce que ça va donner, la mise en perspective de leur travail quotidien, etc. Je ne dis pas que c'est négligé, mais ça pourrait être renforcé bien plus que ça ne l'est, avec des équipes finalement plus tournées vers ces questions-là. Évidemment, ça veut dire des équipes qui sont aussi pleinement conscientes des évolutions techniques et technologiques qui ont lieu. Et quand je parle de ça, ça met en jeu plein de notions connexes, comme le télétravail etc, il y a évidemment dans ce que l'on évoque, ces questions qui peuvent apparaître.

Alors justement, en parlant d'accompagnement au changement, j'aurais voulu aborder la question du web sémantique, puis qu'émerge un nouveau web « 3.0 », articulé autour de la donnée. Déjà, est-ce que vous êtes familier avec ces concepts-là ?

Alors, on n'est pas forcément familiers de ces concepts, ce dont on est familiers, j'allais dire, c'est de la description de la donnée, puisque depuis une dizaine d'année, en tout cas en Gironde c'est prégnant, peut-être que dans d'autres départements ça ne l'est pas du tout, mais nous ici en Gironde on est acculturés, en quelque sorte, à la description de manière structurée, et notamment l'indexation des données. Et on a essayé, évidemment, en réfléchissant à tout cela, de se dire à quoi ça pourrait servir et comment ça pourrait être utilisé. On est familiers de l'intégration de certains de nos projets à des portails nationaux ou internationaux, essentiellement européens, mais effectivement depuis dix ans on est allé s'inscrire dans des opérations type La grande collecte, type « le cloud », etc, qui visaient à la fois à nous apporter de l'argent mais aussi à inscrire nos projets dans des schémas, des opérations, qui visaient au partage de la donnée sur le web sémantique.

Et est-ce que vous connaissez certains projets, des ressources ou des gisements de données qui sont mis en place dans le domaine de la culture et qui reposent sur ce web des données liées ?

Alors, des projets il s'en développe certainement beaucoup, notamment via les ministères. Maintenant, c'est vrai que ça ne redescend pas forcément systématiquement, je parlais des projets type Europeana, des projets comme ce portail qui se développe pour les registres matricules mais lequel sont également inscrites des données relatives aux morts pour la France, etc. Bon, ce sont des projets qui sont inscrits dans des logiques de commémoration, des logiques mémorielles en quelque sorte, mais qui, évidemment, recouvrent, au-delà de cette idée mémorielle, l'idée d'un partage d'information au niveau mondial. Donc, on n'est pas forcément informés de tous les projets qui se développent, mais l'idée c'est de s'inscrire dans ceux qui se développent spécifiquement pour nos services propres. Alors, ensuite, il y a évidemment les systèmes type Wikipédia, sur lesquels on a une information très peu partagée, puisqu'il n'y avait pas trop de monde, mais lors de la journée internationale des archives vendredi dernier, avec une présentation par l'équipe girondine qui anime Wikimédia, des différentes possibilités qui pourraient être les nôtres de s'inscrire dans ces schémas-là. Alors, il y a certains services patrimoniaux, culturels, même archivistiques, je pense aux archives nationales, peut-être à d'autres, je ne les connais pas tous, qui ont voulu conventionner avec des partenaires de ce type et inscrire leur action et le partage d'information, et c'est aussi une forme de publicité du service qui participe à ce genre d'actions, en signant des conventions de partenariat, en faisant venir des wikipédiens sur place, les associant en binôme avec ceux qui s'occupent des fonds d'archives ou de la description des producteurs ou de la description des données et essayer que tout cela puisse coïncider et que ce partage d'information puisse être fait. Mais là, j'allais dire, ce n'est pas quelque chose qui a été systématisé et imposé. On aurait pu dire, au niveau de l'État, puisqu'on a quand même cette double tutelle à la fois territoriale, mais aussi ministérielle, on aurait pu imposer, et on voit que quand certaines opérations sont imposées, c'est quand même plus... enfin je pense à l'opération La grande collecte, qui est une opération venue de l'échelle européenne et qui a essaimé sur la base du volontariat, certes, mais avec de tels moyens en matière de communication et en matière organisationnelle etc, qu'il n'y avait que peu d'intérêt à ne pas y participer, en tout cas dans des services comme le nôtre. On voit bien le succès de ce type d'opération, dès lors qu'elles sont imposées au niveau haut. Si on nous imposait de conventionner... si on imposait à chaque service d'archive départemental ou à chaque département de conventionner par exemple avec Wikipédia, on pourrait développer des actions de manière organisée et harmonisée, bien plus grande, qu'on ne le fait aujourd'hui.

Et comment est-ce que vous percevez les enjeux de ces partenariats ?

Alors, je l'ai un peu dit à l'instant, les enjeux pour une structure comme la notre, c'est de mieux se

faire connaître d'abord, et de mieux faire connaître évidemment... quand je dis mieux se faire connaître, c'est mieux faire connaître ses missions, mieux faire connaître, de manière très pragmatique, les horaires d'ouverture de la salle, le volume des fonds conservés, le volume de la collecte, le volume des fonds traités, mieux faire connaître, finalement, l'organisme, et mieux faire connaître les fonds que cet organisme conserve. Et ça, c'est évidemment vrai pour une structure comme la nôtre, comme c'est vrai pour tout autre type de structure, qu'elle soit privée, publique, associative, ou je ne sais quoi.

Et de votre côté, vous faites une distinction entre les technologies du web sémantique et le concept de web des données, vous personnellement je veux dire ?

Très franchement, non, je laisse cela à mon adjointe qui, elle, s'est familiarisée, je pense, sur le tas, mais qui se familiarise à ce genre de concepts et qui peut ensuite interagir dans les décisions que l'on prend.

Et comment est-ce que vous envisagez...

Et après, on est aussi dans une structure qui dépend d'une structure plus grande, c'est-à-dire qu'on est dans une collectivité, et c'est le rôle d'une direction des systèmes d'information et du numérique que d'impulser, dans une direction comme la nôtre mais aussi dans l'ensemble des directions de la collectivité, il y en a quand même une bonne trentaine, une action qui soit concertée et harmonisée. On ne peut pas faire monter telle ou telle direction en puissance en laissant tomber toutes les autres.

Du coup, pour creuser encore un peu dans cette direction-là, comment est-ce que vous concevez la place du module référentiel dans le SAEM ?

Pour moi, c'est une place essentielle, puisque c'est un peu, je le compare souvent à l'archivistique traditionnelle, mais c'est peut-être dû à mon âge aussi, mais dans l'archivistique traditionnelle, le point d'entrée c'est essentiellement l'outil de recherche, la description des fonds. On peut faire entrer des versements, si vous n'avez pas de bordereau de versement, qui constitue un premier instrument de recherche, vous ne savez pas ce que vous avez fait rentrer et vous ne pouvez absolument pas en informer le public, l'utilisateur, etc. Donc le référentiel, pour moi, c'est le module phare, de tout système, qu'il soit traditionnel ou beaucoup plus innovant.

Du coup, je sais qu'il y a un projet de refonte du site internet en ce moment, comment est-ce que, aujourd'hui, vous percevez les fonctionnalités d'un site web et d'un portail d'archives ?

Alors, ce projet de refonte, il était indispensable, puisque l'on a un site internet assez ancien, une dizaine d'années aussi, donc pour un site internet c'est évidemment ancien. Il y avait une problématique... [le téléphone sonne, l'enregistrement est coupé]

Donc, ce site internet, ancien, pour nous il a une problématique, c'est qu'on avait d'un côté le site institutionnel, avec les infos pratiques, avec les actualités, les manifestations que l'on pouvait organiser, et de l'autre côté il y a GAEL (Gironde Archives En Ligne), qui constitue le cœur, l'essentiel de ce qui est consulté via le site, puisque c'est, vous le savez, l'interface de consultation des fonds via leur instrument de recherche. Donc, l'idée, c'était de ne plus avoir qu'un seul outil, qui intègre la description des fonds et leur consultation, mais aussi tout ce qui peut être mis en ligne et disponible à un internaute, de manière, en dix ans les techniques se sont développées, et maintenant on met évidemment beaucoup plus de formats de fichiers qu'on en mettait il y a une dizaine d'années. Et évidemment, ce temps long pour un site web, n'a pas été un temps statique, c'est-à-dire qu'en dix ans, on a accumulé la mise en ligne de données et on va continuer à le faire puisque l'on a développé des opérations de numérisation, des opérations de traitement et de publication d'instruments de recherche, etc. Donc il y a cette problématique qui est en jeu, là. Alors, moi ce que j'attends du nouveau site, c'est, d'abord, qu'il soit peut-être beaucoup plus ergonomique que ne l'était l'ancien. C'est une réflexion, d'ailleurs, que l'on a eu depuis longtemps, mais on présente les données suivant un cadre de classement des archives départementales qui remonte

quand même à 1840. Donc, en 2020, est-ce qu'il est toujours pertinent de présenter, notamment les fonds que nous conservons, suivant un cadre de classement qui remonte presque à 180 ans. Donc, beaucoup réfléchissent à la perception que pouvait avoir un internaute lambda, alors évidemment ceux qui fréquentent les services d'archives tous les jours connaissent parfaitement notre cadre de classement, et eux peuvent trouver très utile de se référer à cela pour s'orienter dans nos données, mais voilà... Il faut aussi que l'internaute qui ne connaît absolument pas le cadre de classement des archives et qui voudrait simplement avoir une information sur sa famille, sur sa commune, son entreprise, sa société ou sur tel sujet sur lequel il pourrait travailler, son club de sport etc, puisse en trouver, sans tenir compte de ce carcan, qui nous est imposé en termes de gestion, en quelque sorte, mais que ne devrait pas l'être à quelqu'un d'extérieur. Donc, c'est essentiellement ça, que l'on attend, je pense, d'un site internet, aujourd'hui. Certes, on pourrait vous dire : « on attend de pouvoir mettre des vidéos en ligne, sur la conservation préventive, sur la dernière manifestation qui a eu lieu, des expositions virtuelles, etc. » Bon, ça, évidemment ça peut être le moyen de le faire et de faire diffuser, mais on a d'autres moyens, on peut très bien faire un film de notre entretien et le diffuser, le poster via YouTube dès demain et finalement, on valorise à la fois l'institution et nos propres missions. Bon, l'objectif, pour moi, le site internet c'est la fenêtre de la structure que l'on représente et c'est les missions essentielles de cette structure qui doivent être mises en avant. Et comme je le disais tout à l'heure, en parlant du module référentiel, pour moi, la mission essentielle, c'est de permettre la consultation d'une donnée, d'une archive, quelle qu'elle soit, où qu'elle se trouve, etc. par le biais, soit de sa description, soit de sa visualisation, si elle a fait l'objet d'une numérisation. Mais comme vous le savez, la numérisation aujourd'hui, au bout de dix ans d'effort, elle ne concerne à peine qu'un peu plus d'1 % de l'ensemble de nos fonds, et c'est une proportion qui n'ira pas forcément en s'améliorant, dans la mesure où aujourd'hui, on collecte beaucoup plus d'archives qu'on en numérise. Quand on va numériser 300 000 documents, on va en faire entrer, peut-être, plusieurs millions chaque année. Donc cette proportion... il ne faut pas voir dans les sites une manière de consulter virtuellement l'ensemble des fonds que nous conservons, ça non, c'est une façon de les valoriser, de mieux les faire connaître, c'est une façon de permettre à l'internaute de savoir que ça existe, mais pendant longtemps, rien ne remplacera la consultation physique du document, puisque l'on aura que cette possibilité pour le faire.

J'aurais juste voulu revenir sur un dernier point. Vous avez dit que vous étiez tenus de considérer tous les usagers comme égaux, mais j'aurais voulu revenir sur les enjeux de la recherche. Et justement, avec l'évolution du numérique et le mouvement des humanités numériques, comment est-ce que vous percevez ces évolutions ?

Alors, je les perçois évidemment à travers les relations qu'on peut développer avec les universitaires. Il y a quatre ans, quand j'assurais l'intérim de la direction, j'avais passé une convention de partenariat avec l'université de Bordeaux Montaigne, qui est en cours, d'ailleurs, de renouvellement. Mais l'idée, c'était que l'université connaisse mieux et qu'elle puisse s'intégrer pleinement dans nos missions quelles qu'elles soient, aussi bien dans la collecte, puisque l'idée c'était quand même de pouvoir avancer sur la collecte des archives de l'université de Bordeaux Montaigne, mais aussi de pouvoir s'impliquer dans nos missions de traitement intellectuel en faisant venir éventuellement, ce qu'on a fait, pas mal d'étudiants, pour les former un peu au traitement archivistique, mais également nos missions de valorisation par le biais d'expositions partagées, qui peuvent être physiques ou virtuelles. Mais aussi dans l'autre sens, de mieux connaître les activités, les actions de l'Université, notamment dans le domaine que vous évoquez. Aujourd'hui, il me semble que ce n'est pas pleinement abouti, c'est-à-dire que souvent on est contactés, mais comme pourvoyeur de fonds, de documents, qui peuvent rencontrer des champs d'étude dans telle ou telle discipline, et là on nous dit : « on travaille sur un projet de virtualisation de l'œuvre de Montaigne, ou de Mauriac etc, on sait que vous avez tel ou tel document qui les concernent, est-ce que vous pourriez les numériser et, en quelque sorte, est-ce qu'on pourrait ensuite les intégrer dans nos

projets ? ». Donc c'est plus cette relation-là qu'on a développée aujourd'hui, dans ce type de programmes, qu'une relation d'acteurs pleine et entière. Mais là encore, je pense qu'il y a une évolution possible et souhaitable, que des services comme le nôtre puissent avoir une action plus importante.

Annexe 7b : transcription de l'entretien avec Delphine Jamet

Entretien avec Delphine JAMET, adjointe du responsable du service collecte et traitement aux Archives départementales de la Gironde et membre du COPROJ.

Réalisé dans son bureau, en fin de matinée, le 14 juin 2017.

Donc, déjà est-ce que tu peux commencer par me parler un peu de toi et de ton parcours ?

Alors, mon nom est Delphine Jamet, je suis attachée de conservation du patrimoine, titulaire depuis peu. Je suis adjointe au chef du service de la Collecte et du traitement et j'ai commencé, enfin j'ai fait la faculté d'Histoire, ça fait 15 ans que je suis aux Archives, j'ai un DEA, et j'ai commencé par traiter les fonds de la Seconde guerre mondiale. Et donc en traitant les fonds de la seconde guerre mondiale je faisais du traitement et de la recherche et j'ai très vite commencé à saisir dans le module de... notre outil de... notre outil qui nous permet de faire les instruments de recherche électroniques et après je me suis occupée de la rétro-conversion des instruments de recherche, ce qui m'a amené à l'informatique, enfin à lier informatique et archives.

Ok... Et donc, ça t'amène à interagir avec quels types de professionnels, à la fois dans le service et pour les projets ?

Alors tout ce qui a été rétro-conversion m'a amenée à travailler avec des prestataires extérieurs puisqu'on passait par des marchés publics. Euh, mais aussi du coup, parce qu'il y a eu beaucoup, dans ces marchés de rétro-conversion, il y a eu des marchés de rétro-conversion pure, on partait du papier et on allait à la base de données mais aussi après on a dû faire de la normalisation, pour respecter les normes archivistiques, et notamment de l'indexation. Donc là, j'ai été amenée à travailler avec le Service interministériel des Archives de France sur les questions d'indexation, pour bien comprendre les normes, suivre les normes, etc. Euh, après il y a eu, qui d'autre ? Les collègues d'autres services qui étaient... qui travaillent déjà sur ces questions-là pour en discuter, c'est toujours bien d'avoir un avis extérieur. Et après effectivement sur les marchés numérisation, il y avait à l'époque Pascal Romain qui était aux Archives départementales et qui est après parti à la DSI mais bon le lien s'est fait aussi avec la DSI de façon plus approfondie pour voir comment on rétro-convertissait tout ça et tout ça est aussi lié à la numérisation des archives, en fait. Donc, pour que tout soit cohérent, donc là on a essayé de travailler comme ça. Parce que je ne t'ai pas dit que j'avais été chef de projet pour l'outil de mise en ligne, Pléade, pour GAEL. Donc là, pareil, j'ai été amenée à travailler de façon étroite avec la DSI, et les prestataires bien sûr, qui développent. Et pareil, sur l'outil Arkheïa j'ai été amenée à travailler avec Louis Colombani de façon assez proche pour faire faire du développement, pour mieux comprendre l'outil, pour voir ce qui était possible...

Du coup ça fait combien de temps que tu travailles sur le projet de SAEM ? Depuis le début ?

Non, pas depuis le début. Non, en fait au début c'était Véronique Pisani qui travaillait sur le projet de SAEM, et donc moi j'ai intégré le groupe projet je crois en décembre 2012 ou quelque chose comme ça. Donc ça fait cinq ans quand même *rires*

****Rires* Et... comment est-ce que tu as été amenée à travailler sur ce projet ?***

Alors, moi j'ai été amenée pour faire un binôme avec Véronique [Pisani], qui était à l'époque chef de projet, et effectivement moi c'est plutôt tout mon côté spécialiste des métadonnées, des instruments de recherche, etc. qui m'a amenée dans le projet. Parce que le projet était très "collecte" au départ, finalement c'est pas... moi je suis, je comprends... je fais un peu plus de collecte depuis que je suis adjointe au chef de service mais avant j'en faisais pas du tout donc j'avais pas forcément cette vision-là, de records management, de collecte, etc. Et donc moi j'étais plutôt sur la pérennité, les métadonnées, comment on décrit les archives définitives, etc. Donc c'est plutôt pour ça que je

suis arrivée, et pour essayer d'apporter cette vision en plus en fait.

Ok, et du coup, ça te... ça t'amène à échanger et à collaborer avec qui principalement, là aussi ?

Pour le système d'archivage électronique ? Avec tous mes collègues du COPROJ, bien entendu, mes collègues ici en interne : direction, chefs de service mais aussi les autres directions, notamment la mission d'action territoriale, par rapport aux communes et aux archives des autres territoires. Le SIAF, bien entendu. Euh, d'autres collègues d'autres départements qui s'occupent de systèmes d'archivage électroniques, d'archives contemporaines de façon plus générale. Hum, avec les prestataires aussi avec lesquels on travaille et... je pense que j'ai fait le tour là.

Et du coup, les échanges notamment avec les informaticiens... Comment est-ce que t'as perçu ça après ton arrivée dans le projet ? Parce que...

Alors, je pense qu'on a eu des informaticiens qui ont très vite compris le métier des archives. Vraiment, enfin... Je pense que... L'organisation du COPROJ a été essentielle et je pense qu'on a eu de très bons interlocuteurs, parce que même si on a changé d'interlocuteurs au cours du projet, à chaque fois les informaticiens se sont mis à essayer de bien comprendre les archives et nos problématiques, avec toujours une écoute réciproque, pour que le côté technique ne l'emporte pas sur le côté fonctionnel, forcément tout le temps. Même si effectivement chacun défend un peu son bout de gras, et c'est normal, c'est de bonne guerre. Mais il y a toujours une très bonne entente, de tous les côtés, sur ça. Et c'est ce qui fait aussi la force du projet, je pense, parce qu'ils comprennent, et nous on comprend mieux aussi le côté technique. On a fait l'effort de comprendre un peu mieux même s'il y a des choses où on n'a pas envie de le faire non plus parce que c'est pas notre domaine. Eux ont vite compris le monde des archives, et l'informatique en fait... le monde des archives est tellement formalisé que ça rentre bien, finalement, dans l'informatique. Donc, non je pense que c'est... ils se sont tous formés avec nous, ça a été un bon... Et après, on a eu la chance d'avoir Pascal Romain, qui avait la double casquette, d'attaché de conservation et en même temps d'ingénieur. Il a quand même les deux concours.

Donc, tu dirais qu'il y a quand même un certain équilibre au niveau des forces en présence ?

Alors, oui, je dirais que oui, même si c'est plutôt la DSI qui gère tout ce côté administratif et tout donc c'est eux qui font la plus grosse part de travail sur ce genre de choses. Après en tout cas pour faire valoir les visions, les objectifs, il y a un vrai rapport d'équilibre.

Et justement, cet équilibre-là, il n'a pas évolué tout au long du projet ? Il y a toujours eu cette volonté de... de donner la parole un peu à chacun ?

Il y a toujours eu cette volonté mais effectivement je pense qu'aux débuts, peut-être que les archivistes étaient plus... plus présents on va dire mais... parce que... dès que les informaticiens montent en compétence côté archives, en fait finalement, nous on a pas besoin de... en fait ils comprennent vraiment les mécanismes donc finalement le côté fonctionnel, ils finissent par le comprendre. Après, c'est vrai qu'on a tendance peut-être à se relâcher un peu de côté-là, parce qu'on sait qu'ils ont compris, donc une fois qu'on a exprimé une fois ou deux le besoin, on a pas besoin de répéter dix fois donc pour ça c'est vrai qu'ils peuvent nous remplacer... enfin, pas nous remplacer mais en tout cas se substituer à nous, et à certaines réunions on n'est pas obligé d'être tous là.

D'accord, et...

Mais là, c'est surtout qu'on est en phase de production, et là effectivement ils prennent plus de poids parce qu'on est vraiment dans la phase plus technique, tout ce qui est mise en production, stabilisation des outils, etc. Donc là, il y a quand même un poids plus technique, moi je trouve en ce moment, que ce qu'il y a eu avant. Par exemple pour les sprint plannings, avant on réfléchissait beaucoup plus, on faisait des ateliers, là on n'a pas fait un atelier depuis 6 mois, quoi. Donc,

typiquement là voilà, on voit que eux prennent le dessus là-dessus mais, ce qui après paraît légitime aussi, et normal dans ce type de phase.

Donc, finalement, d'après toi, les différentes cultures professionnelles, les différentes visions... elles ne compliquent pas tant que ça les échanges ?

Non, parce que déjà je pense qu'on est dans une équipe où il y a des gens intelligents *rires*. Non, mais après c'est, je sais pas comment dire... ça dépend beaucoup des personnalités aussi, donc... Et il y a une vraie volonté de travailler en commun. Donc si on a cet état d'esprit là, ça marche, si effectivement on n'a pas cet état d'esprit, ça ne marche pas.

Et au niveau de l'archivage électronique, dans son ensemble, est-ce que tu dirais qu'il y a une grosse modification des pratiques de l'archivage "traditionnel", on va dire, avec justement l'arrivée de l'archivage électronique, de la mise en ligne des documents sur les portails des services d'archives, ce genre de choses ?

Alors, de toute façon, la grosse mutation s'est faite avec la mise en ligne des documents numérisés, déjà, et la rétro-conversion des instruments de recherche, et l'informatisation de tout ça. Là, déjà il y a un gros travail de fait, après l'archivage électronique... l'intérêt c'est qu'il faudrait qu'on soit plus au cœur du système de production du document. C'est là, en termes d'organisation, que je pense qu'on pêche un peu, parce qu'on n'est pas assez au cœur du système, pour la production du document, pour archiver à terme. Et c'est en ça effectivement que nos pratiques devraient être plus modifiées. En fait, nos tableaux de gestion devraient être tous mis à jour avec la production documentaire nativement numérique et en plus, du coup, donner des DUA, les métadonnées qu'il faut, les types de documents, des documents types, ce genre de choses. On devrait être au cœur de ça et notamment pour la GED courante. C'est là que... autant, je pense que sur le système d'archivage électronique on a bien travaillé en commun, autant sur tout ce qui est mise en place de la GED, pas assez, sauf que c'est lié à terme... Et pour la mise en ligne...

Et du coup, si c'est pas les archivistes, qui gère un peu ces aspects-là ?

Là, pour le coup, ça a été plutôt la DSI.

C'est la DSI ? D'accord.

Oui, après ils nous ont toujours associés, mais typiquement nous on avait aussi demandé à être pilotes pour la mise en place de la GED transverse, pour la direction, parce qu'il valait mieux qu'on sache il fallait l'utiliser pour nous-mêmes, pour après pouvoir... Et on ne l'a pas été, donc voilà, là je pense qu'il y aurait dû y avoir un binôme vraiment bien défini, comme il y a aujourd'hui pour le COPROJ, au niveau du SAE. Pour la GED transverse, il aurait fallu un binôme bien défini aussi, avec une organisation type COPROJ aussi, en fait, [avec des réunions] toutes les semaines, etc., ce qui n'a pas été. Donc là, c'est sur ce côté-là que ça pêche aujourd'hui à mon avis pour la mise en œuvre de ça et pour l'instant on fait des flux pilote etc pour l'archivage électronique, mais il va y avoir, bien entendu, cette production documentaire qu'il va falloir archiver aussi à terme, donc là je pense que c'est ça qui... Après, pour la mise en ligne, je pense que ça va être la même chose donc les documents numérisés, enfin il n'y a pas de... pour moi ça ne pose pas de soucis en termes de... en termes de changement, de conduite de changement, c'est pas ça qui va être trop problématique. Par contre, là où on a beaucoup évolué aussi, c'est le fait de contextualiser la production. Donc avec le fait qu'on aille sur de l'archivage électronique, on a été obligé de mieux contextualiser la production documentaire et donc du coup ça a impacté nos pratiques sur le papier. Parce que si on le fait pour le numérique, alors pourquoi ne pas le faire pour le papier, donc on a mis en œuvre des procédures pour pouvoir faire notamment les notices d'autorité.

D'accord.

Et l'indexation au niveau haut des versements, ce genre de choses...

Hm. Et justement ces grands principes de l'archivage, quand ils sont appliqués à l'environnement numérique, est-ce que tu dirais que ces nouveaux standards qui s'imposent, je pense au modèle OAIS, aux normes 30300... Est-ce que ça coïncide avec les pratiques actuelles ou... ?

En gros, oui.

Oui ? Il n'y a pas de "clash" entre...

Non, en fait ça se colle assez, enfin... finalement on voit... enfin, quand on regarde en tout cas le modèle OAIS, quand on le regarde c'est un peu ça : les documents qu'on prépare, à verser, après les documents versés, les documents communiqués, finalement c'est la chaîne archivistique, avec la planification, la pérennisation et la préservation. Pour moi, ça rentre complètement, quand on regarde le schéma c'est ce qu'on doit faire aussi dans le papier. Parce qu'ici, c'est ça, on a de la collecte, on a du traitement, on rentre les métadonnées, on a de la conservation préventive donc la pérennisation. Et on a de la planification, on a de la communication, on a tout. Donc à mon avis, on a tout ce qu'il faut en conformité avec le modèle OAIS. Donc ça n'a pas vraiment révolutionné tout ça, après la problématique c'est de s'approprier le langage, le vocabulaire, qui est plus complexe. Parce qu'avant, on faisait les choses de manière plus ou moins instinctives et là on nous l'a normé, on nous l'a normalisé *rires*. Donc il faut s'approprier ça mais finalement quand on est assez pragmatique, on voit que c'est la même chose. Donc c'est pas... je dirais pas que c'est forcément très complexe.

Et est-ce que tu dirais qu'on pourrait aujourd'hui parler d'une sorte de convergence des approches entre ce que si fait dans ce qu'on appelle le records management, l'archivage traditionnel et justement la vision qu'ont les informaticiens de l'archivage.

Répète la question...

Est-ce que tu penses qu'il y a une convergence des approches, un objectif commun et que...

L'objectif commun c'est de conserver et de communiquer.

Oui, oui, mais... les informaticiens ont peut-être une vision axée plus autour de la sécurité, par exemple...

Alors nous c'est une vision qu'on a quand même parce qu'on est obligé de l'avoir et...

Bien sûr, bien sûr, oui...

Donc ça c'est quelque chose qui, effectivement, nous a... Enfin je pense que travailler sur l'archivage électronique, c'est la valeur probante qui a été rajoutée en plus. Dans le papier, on l'avait naturellement, on avait le document signé, on ne s'en occupe pas trop, forcément. Tandis que là, là moi je pense que c'est ce domaine-là qui a changé pas mal, et dans le groupe projet avant on avait Aude Banabera, qui elle était plutôt juriste, et elle, elle a beaucoup amené et ça a permis de faire évoluer nos visions, nous en tant qu'archivistes classiques, en fait je pense. Et donc ça, ça m'avait beaucoup apporté, je me rappelle, à l'époque, parce que c'était quelque chose que je méconnaissais totalement. Alors je sais qu'on a le Code du patrimoine, les textes et règlements etc, mais finalement... Et après, par rapport au records management, ce n'est quand même pas la même chose. Entre les archives définitives, qu'on fait verser ici et le records management, c'est pas tout à fait... Alors l'archivage électronique, oui, mais l'archivage papier en tout cas on traite plutôt les archives définitives, donc c'est plus de la collecte d'archives définitives que du records management. Le records management, on va le faire pour le Département et c'est tout, au final. Donc effectivement, avec les archives électroniques, on est obligé de voir cette dimension-là, et de voir les archives courantes, intermédiaires et définitives, parce que finalement, la limite, elle n'y est plus, quoi, elle y est très peu. Donc c'est vrai que là, là ça change pas mal.

Ok, est-ce que toi tu as suivi des formations par rapport à tout ce qui est archivage électronique ?

Ouais, j'ai fait une formation, donc sur les métadonnées, etc. Et j'ai fait la formation du SIAF à

l'époque, c'était une semaine de découverte de l'archivage électronique où on avait des formations théoriques, et en même temps on allait dans des services qui avaient déjà mis en place des SAE, comme La Poste, L'INPI, qui en avaient fait, Orange, du coup c'était assez intéressant.. Voilà, j'ai fait ça, qu'est-ce que j'ai fait d'autre comme formation...

Hm, sinon... bon je passe un peu du coq à l'âne là mais les instruments de recherche, dans leur forme actuelle, est-ce que tu penses que c'est toujours des outils pertinents, à la fois pour les archivistes qui les réalisent et pour les publics ?

C'est compliqué en fait. En fait, le problème c'est la hiérarchisation, sauf que effectivement avec ce côté hiérarchique, enfin c'est... on ne répète pas l'information supérieure, on ne la répète pas et on la met au niveau supérieur. Moi je trouve que ça, oui... Après c'est la façon dont on publie ces instruments de recherche qui n'est pas forcément pertinente, de les décrire comme ça, moi ça... Enfin d'abord on est habitué, c'est une pratique qui est bien ancrée depuis les années 90, donc ça va faire presque 30 ans que c'est comme ça, même plus, je ne sais plus trop de quand date ISAD(G) mais voilà, c'est quand même une pratique bien ancrée donc c'est compliqué... et qui est normée, qui est claire aussi... Donc moi ça me semble assez pertinent pour les archivistes, par contre c'est la façon dont on les exploite effectivement, qui peut ne pas être pertinente ni pour les archivistes d'ailleurs, ni pour le public. Et effectivement complexe, l'approche, etc. Donc simplifier... et d'éclater tout ça peut être effectivement une bonne solution.

Du coup, pour passer sur la partie web sémantique, je voulais savoir ce que toi tu connais de ces technologies-là, et comment tu as pu te familiariser avec ces concepts...

Alors, c'est Pascal Romain qui m'en a parlé en premier, donc effectivement, qui a une cette vision-là... Donc il m'a expliqué, il m'a expliqué souvent. Je me rappelle, une fois, dans une conversation, il m'a dit « c'est normal, Delphine, que tu ne comprennes pas tout, quand on a des notions neuves il faut au moins les répéter sept fois avant de les intégrer, surtout quand elles sont un peu complexes ». Donc... et après, après de toute façon il y a eu la feuille de route aussi du ministère de la Culture, qui à mon avis est un vrai document de référence, donc... parce qu'une fois qu'on commence... donc Pascal m'en a parlé comme ça un peu en "off" etc, alors voilà... Du coup on se renseigne, il y a cette feuille de route, qui est importante, on en a discuté aussi du coup avec pas mal de collègues des archives nationales, du SIAF et de la BNF, donc voilà, ça c'était important. Ça paraît logique aussi, c'est très logique tous ces triplets, le fait que tout rebondisse avec le web, etc., moi ça me parle aussi. Après, il y a eu, et j'avais beaucoup aimé, bon c'est pas du web sémantique mais c'est un peu la même approche quand même au final, c'est un projet sur les notices producteurs... Comment il s'appelait ce projet... C'est un projet américain, d'une université américaine, qui agrège les notices d'autorité et ils en font des graphes, du coup il y a des graphes et si on met La Fayette par exemple, du coup on va voir tous les... toutes les personnes en relation avec La Fayette, toutes les personnes avec, en relation avec... enfin, toutes les relations du style où sont conservées les archives, les documentations, etc. Donc ça c'est un outil qui est plutôt pas mal, qui m'avait bien parlé au final, qui m'a permis d'avoir une vision de ce qu'on pourrait faire finalement avec nos instruments de recherche. Après, effectivement, et au fur et à mesure du temps qui passe, il y a l'ontologie RiC qui arrive, et qui là par contre est une vraie révolution dans la manière de décrire les archives. Donc même si on continue à le faire en ISAD(G), finalement c'est les outils qui derrière feront du... de l'ontologie. Il n'empêche que le modèle conceptuel est... permet... va quand même pas mal révolutionner les choses, même si on fait déjà plus ou moins ça.

D'accord...

Mais c'est de segmenter tout comme ça, qui va...

Alors, pour toi, il n'y a pas une réelle opposition entre ce qui se fait maintenant et...

Non, en fait, parce que les éléments qu'on met, nous, par exemple déjà dans ISAD(G) ou dans l'EAD, c'est des éléments qu'on va retrouver dans RiC. Donc, enfin, qu'on le mette à ce niveau-là

ou à un autre niveau, enfin... C'est juste... pour moi c'est pas problématique.

Donc, les éléments seraient les mêmes, mais agencés d'une autre manière ?

Donc ouais, moi ça me paraît pas trop problématique.

C'est surtout quand on a fait des instruments de recherche qui sont déjà bien détaillés, qu'on a déjà bien segmenté les informations. Il y en a qui ont fait des instruments de recherche où ils vont mettre le nombre de pièces dans la présentation du contenu, des trucs comme ça. Nous, on l'a bien mis dans la zone de description physique, les lieux de conservation on les a bien renseignés, enfin... On a bien renseigné zone par zone donc, du coup, finalement champ par champ de l'EAD, donc on va pouvoir redispacher ça correctement dans RiC, au final.

Donc en termes de pratiques, finalement, pour les archivistes... Ça ne change pas...

Moi je pense que ça ne va pas changer grand-chose. Je pense que c'est plutôt une façon de réfléchir à la description, qui va changer, c'est-à-dire que ça va nous permettre... déjà, je crois que ça simplifie certaines choses, ça nous oblige à nous dire « Ah bah lui s'il a un truc, enfin un nom, ça veut dire qu'il peut aussi avoir une date de naissance, etc. » Enfin voilà, ça nous oblige à mettre de plus de... à nous poser plus de questions sur quelle description on veut mettre en fait. Mais sinon, non moi je trouve que c'est... Par contre, ce qui est super intéressant après, c'est de pouvoir faire les liens avec les autres, parce que là où on pêche c'est sur les sources complémentaires, on a jamais de liens pérennes, ce genre de trucs. Donc c'est plus ça qui est intéressant par exemple dans RiC, le fait d'avoir des identifiants pérennes, et ça va pousser par exemple... Moi par exemple quand j'ai un lien vers un instrument de recherche des archives nationales, j'ai un pauvre lien vers leur site où il n'y a pas d'identifiant ARK, donc quand ils changent de système, tout mon système plante, quoi. Donc ça c'est... enfin, il faut revoir tous mes liens, donc effectivement c'est pas le but à terme de revoir à chaque fois ses liens. Donc ça, sur les liens comme ça, je trouve que c'est beaucoup plus pérenne et ça nous permet de faire moins de travail, en fait, de reprise.

Du coup, justement, comment est-ce que tu envisages la place du module référentiel dans le projet SAEM ? Pour toi, c'est quoi son enjeu, sa place etc ?

Pour moi, son enjeu, alors je pense que c'est là où on diverge un peu avec Pascal, quand même. Pour moi, les premiers qui vont l'utiliser, ça va être les archivistes, ça ne va pas être les services. Donc c'est pour ça que pour moi il est important, typiquement pour le thésaurus géographique, qui soit vraiment avec nos formes autorisées du nom à nous, pour qu'on s'y retrouve... enfin que nos formes autorisées soient alignées sur le reste, OK, mais que le point de départ ne soit pas forcément le reste en fait. Ça, je pense que c'est une divergence qu'on peut avoir niveau SI, enfin entre SI et archivistes, parce qu'effectivement ce ne sont pas les services qui vont tout de suite l'utiliser, ça moi j'en suis persuadée. Donc, déjà si nous on l'utilise bien, on pourra bien le faire utiliser ailleurs. Après, moi l'intérêt que j'y vois, effectivement, c'est que les référentiels, moi j'en ai plein, et là je vais en avoir un seul, un seul outil qui permet de gérer tout ça, et donc d'éviter de toujours répéter les choses, d'avoir tout en un seul endroit, que quand on cherche un mot, on sait que c'est là qu'on va le trouver et pas ailleurs. C'est en termes de gestion des données de référence... Nous, on en gère beaucoup, de noms, entre les noms de producteurs, les noms géographiques, les noms de personnes, les noms de notaires, les mots matière etc, on en gère énormément. Donc que tout soit au même endroit, c'est quand même... et souple ! Pour la mise à jour, tout ça, pour moi c'est ça qui est le plus important. Et l'usage aussi... mais c'est là qu'il faut que ça soit bien couplé avec notre système d'information archivistique.

[le téléphone sonne, l'enregistrement est coupé]

Donc oui, pour moi [le référentiel] peut vraiment être pivot, et après ça va sûrement être pivot pour la GED transverse, ça va être intéressant pour, effectivement, pour qu'on ait les mêmes termes d'indexation, en tout cas qu'il y ait des alignements, et du coup que nous on aura beaucoup moins

de travail à faire à terme sur les archives électroniques qu'on va conserver. Donc ça, ça me semble primordial. Nous, on avait besoin d'un outil qui nous permette de faire des notices d'autorité propres, donc là maintenant on l'a donc on est content. Euh, et puis même pour la gestion du SAE... Enfin, le but c'était ça, de ne pas répéter les informations partout, donc à mon avis ça a une place assez pivot.

Et juste comme ça, est-ce que toi tu fais la distinction entre les technologies du web sémantique et le concept du web des données liées et ouvertes ?

sourir Je...

Est-ce que c'est quelque chose qui te parle ou est-ce que...

Ouais, si tu veux, j'arrive pas à l'expliquer, quoi... J'arrive à comprendre, en gros, mais je peux pas expliquer, c'est compliqué pour moi encore...

Et du coup, sur ces aspects, tu n'as pas eu de formation par contre ?

On a eu une formation au web sémantique là, avec Logilab, et sur le web, etc. Mais après, tout le monde parle de web sémantique et de web des données pour la même chose quoi... Donc effectivement en termes de... c'est compliqué. *rires* C'est compliqué de savoir qui a dit quoi, quand et c'est souvent, même dans les présentations et tout, c'est étonnant... Mais je pense que le côté web sémantique c'est plutôt le côté technique, et web des données c'est l'ouverture, ce que ça va donner, quoi.

Et là encore, comment est-ce que tu vois la... Comment est-ce que tu conçois la compatibilité entre ces technologies-là et la description en archivistique ? Est-ce que pour toi, ça coïncide, du coup, avec le modèle RiC ?

Bah oui, moi j'en suis persuadée, ça va bien, ça va faire plein de rebonds. Mais en plus, les archives c'est quand même le but, c'est que comme on est par fonds et par producteur, le but c'est de faire des rebonds tout le temps, puisqu'on va forcément retrouver des choses ailleurs, donc moi je trouve que c'est totalement approprié à ça. Et le modèle RiC a l'air d'aller dans ce sens-là, alors il ne va pas être parfait au début, mais il a bien l'air d'aller dans ce sens.

Est-ce que tu connais certains projets, dans le domaine culturel, qui justement, reposent sur toutes ces technologies, et qui ont déjà abouti...

Huma-Num, ou quelque chose comme ça, comment ça s'appelle... Isidore, leur portail et tout ça.

Du coup, il y a le projet du SIAF, AN et BNF, où ils sont en train de faire le prototype là, avec Logilab, où on a donné des éléments à nous d'ailleurs, pour pouvoir les rentrer dedans. Et du coup il y a aussi le truc [elle lance une recherche en ligne]... Comment ça s'appelle... Excuse-moi... Après s'il y a eu une présentation par le Ministère de la culture. Après, il n'y en a pas beaucoup, je crois, dans le domaine culturel.

Mais en fait, si je te demande ça, c'est surtout pour savoir si tu perçois un peu les enjeux pour les usagers, pour...

Alors, moi ce que je vois c'est qu'en tout cas, pour la recherche à facettes ça facilite vraiment la donne, pour les usagers. Parce que quand tu regardes Bach d'Anaphore, c'est basé sur une ontologie, etc. Et finalement tu n'as plus de formulaire de recherche avancée ou particulier, c'est un formulaire... enfin c'est de la recherche simple, et derrière tu as de la facette en fait. Donc je pense que pour le lecteur c'est plus clair. Mais bon après c'est un avis très personnel, parce que si tu interrogés Georges Cueur, tiens tu devrais peut-être interroger Georges Cueur, qui a un avis complètement différent sur le web sémantique, et sur justement la vision pour la mise en ligne, etc. Tu devrais peut-être l'interroger...

Et pour revenir un peu sur les instruments de recherche, le fait qu'on semble passer du

document à la donnée, qui deviendrait la nouvelle unité de traitement, est-ce que, pour toi finalement, est-ce que ça n'a pas un impact sur la conception de ces outils, qui restent plutôt du domaine du document... C'est-à-dire un document fini...

Tu entends quoi par document et donnée là ? Parce que du coup, ça dépend...

C'est la granularité qui change, les données, c'est ce qui va composer un document.

Pas forcément...

Non, pas forcément, mais...

Non, parce que pour moi un document, que ça soit une pièce, un groupe de pièces, un dossier, c'est plutôt là où est la granularité pour moi. Après, en termes d'archivage électronique, on va avoir les données, enfin c'est plutôt... Nous, de toute façon c'est pas la donnée en elle-même mais plutôt les métadonnées sur les documents, pour le coup, qu'on fait. Donc, oui, on va être obligé d'aller à la pièce, enfin plus je veux dire, mais ça se fera automatiquement en fait, d'où l'importance d'être en amont de la chaîne pour ne pas avoir à les remplir, pour dire que c'est un rapport, un compte-rendu, un procès verbal etc, que ce soit des choses...

Mais du coup, en termes de faisabilité, aller jusqu'à la pièce, c'est quand même, enfin je veux dire...

Mais on ira pas tous... pas pour tout à la pièce, je pense... Enfin, de toute façon on va perdre plein de choses, moi je n'y crois pas du tout...

****Rires* Ok...***

Enfin, ouais, je pense qu'on va perdre beaucoup de choses, on ne fera pas l'histoire comme on a fait l'histoire jusqu'à présent...

Et du coup, vu qu'il y a ce projet de refonte du site et du portail des archives de Gironde, comment est-ce que tu perçois les fonctionnalités d'un site web d'archives aujourd'hui ? Est-ce que c'est toujours le même rôle, est-ce que...

Alors, déjà, je pense qu'il faut qu'on puisse interroger autant les archives papier que les, enfin, les instruments de recherche qui ne renvoient que vers des archives papier, qu'on doit commander en salle de lecture, que les instruments de recherche qui renvoient vers des documents numérisés, et les instruments de recherche d'archives électroniques. Donc ça, faut effectivement faire ça. Après, dans un service d'archives, on a une bibliothèque, donc il faut avoir des résultats qui nous ramènent le fonds de la bibliothèque, on va avoir différentes typologies d'instruments de recherche, qu'on a déjà, mais notamment on va faire de plus en plus d'états de versement, ce qui n'est pas au même niveau qu'un instrument de recherche numérique, etc. Donc mettre tout ça en corrélation, effectivement, est assez complexe. D'où le fait d'avoir, effectivement, une recherche très simple, et pouvoir filtrer après par facettes. Après, en termes d'approche, ça dépend de qui on est et de ce qu'on cherche. C'est-à-dire que quand on est un généalogiste, il faut la recherche la plus simple possible et la visualisation la plus simple possible de ce qu'on veut, et avoir son résultat. Après, quand on est un historien, il faut avoir nos documents bien remis dans le contexte, parce qu'on ne fait pas la même analyse du document s'il n'est pas remis dans son contexte, d'où l'importance des notices producteurs. Mais franchement, le généalogiste, il n'en a rien à faire que ça vienne du tribunal d'instance donc... Nous, on le mettra, mais après ce n'est pas un élément qui doit ressortir. Après, encore, en même temps, pour l'état civil, on a les fonds du tribunal d'instance, des communes, on a les documents qui sont déposés et pas déposés, des communes. Donc quand même il faut bien qu'on fasse ça, qu'il puisse avoir cette information-là à un moment donné, s'il en a besoin. Donc je pense qu'effectivement, le mieux ce serait qu'on se connecte avec un certain profil, et on a des interfaces particulières. Ça, je pense qu'à terme, ça serait bien, parce que...

Donc il faudrait vraiment prendre en compte les différentes typologies de publics ?

Ouais, je pense, mais à long terme, mais... Parce qu'en fait, on voit, on a plein de gens qui ne comprennent pas comment ça marche, même en interne. Moi, typiquement, je ne fais que peu de recherche, je passe directement par le cadre de classement parce que je sais où je vais aller trouver

mon info, et c'est à l'intérieur de mon instrument de recherche que je vais faire ma recherche. Là par exemple, je t'ai montré il n'y a pas très longtemps. Et il y a plein de gens qui ne savaient pas, en interne, qu'on pouvait chercher dans un seul instrument de recherche. Donc, voilà, ça c'est un peu...

D'où ma question, tout à l'heure, sur les instruments de recherche...

Mais oui, mais après le problème c'est qu'il y a différents niveaux de connaissance sur tout ça, quoi, donc... Après, je pense que l'indexation est primordiale, si on veut éviter le bruit, la recherche plein texte, il faut bien indexer quand même, un minimum. Et déjà, si chaque instrument de recherche pouvait afficher le contexte historique, déjà ça permet de limiter pas mal de choses, en tout cas pour la recherche historique, me semble-t-il. Parce que si on tape "guerre", si on ne tape que "guerre", on va avoir trois millions de réponses alors que si c'est juste la seconde, autant pointer juste sur la seconde. Après, je vois surtout des choses qui se mettent un peu à jour toutes seules, on publie des métadonnées, que le site internet va chercher et hop ça se met à jour, ça serait bien, ça. Et surtout, que tout converge, quoi, je pense que le but est quand même là, quand on va chercher une délibération sur un sujet, par exemple l'eau, que ça aille chercher la gestion de l'eau du 19e jusqu'à aujourd'hui, toutes les délibérations qui ont traité de la gestion de l'eau, ça peut être pas mal. Donc, pour ça, il faut avoir une recherche qui doit aller pointer sur les instruments de recherche rétro-convertis et dans l'archivage électronique. Parce qu'il y a des fonctions qui traversent le temps, quand même...

Et justement, pour en revenir aux enjeux en termes de recherche, quel impact est-ce que tu penses que ça va avoir, tout ce... la mise en ligne de corpus d'archives numériques. Je pense au mouvement des humanités numériques, ce genre de choses. Est-ce que, toi, tu as réfléchi à ces questions-là de ton côté ?

Alors, pour moi, c'est pas forcément les humanités numériques, c'est plutôt l'accès aux documents administratifs en fait.

D'accord.

Je pense qu'avec l'archivage électronique, on va capter plus vite les documents, et donc certains documents vont être plus vite mis en ligne, et ça peut avoir un lien avec l'open data assez rapidement, je pense. Même s'il y a des jeux de données qui sont souvent mis à jour dans les administrations etc, il n'empêche qu'une étude sur l'eau, les délibérations, ce genre de choses... ça permet une plus grande transparence de la production administrative, à mon avis. Et donc, ça, ça peut être pas mal. Après, il va y avoir la complexité de la communicabilité du document, donc l'identification, tous ces enjeux d'identification et de sécurité sont importants, puisqu'il y a des documents qui sont soumis à dérogation, d'autres non, etc. Mais en tout cas, ça va permettre aux gens de savoir ce que produit l'administration, et ça je trouve ça assez intéressant parce que finalement, il y a une vraie méconnaissance citoyenne de l'administration et la mise en ligne des métadonnées, de dire « bah ça, ça existe », que vous ayez accès ou pas au document mais « ça, ça existe », je pense que c'est une vraie transparence. Plus que... parce que pour l'instant, l'archivage électronique, ce n'est pas forcément des documents d'histoire, enfin, c'est vraiment de la production courante, enfin récente. Donc on n'a pas forcément le recul pour faire de l'histoire avec... ça dépend des documents, bien entendu, mais à la limite les documents intéressants pour faire de l'histoire, ils ne sont pas librement communicables tout de suite aussi. Par contre, pour tout ce qui est ouverture des données administratives, là je pense qu'il y a plus de compatibilité et d'impact immédiat.

Et au niveau de... pas au niveau de la description archivistique en elle-même, mais au niveau des grands principes qui régissent l'archivage, comme le principe de respect des fonds, des choses comme ça, est-ce que là aussi tu crois que finalement l'impact est négligeable ?

Non, ça nous oblige à être beaucoup plus consciencieux et à contextualiser dès le départ, ça c'est sûr. Mais ça, on le savait, par défaut, que c'est tel producteur qui avait fait tel versement, et tout...

Et encore, moi je vois parfois, sur les archives papier, qu'il y a des erreurs entre service versant, service producteur. Là, par contre, ça nous demande d'être vraiment exigeants avec nous-mêmes sur ça, en fait. Parce qu'effectivement, si on veut reconstituer les documents produits, enfin la production documentaire d'un service, il faut absolument que ces métadonnées y soient, donc ça c'est un des points fondamentaux. Parce que la contextualisation, elle est hyper importante pour le respect du fonds. Le principe de base du respect du fonds, ce n'est pas pour embêter les gens, c'est pour faire l'histoire. Effectivement, on n'a pas les mêmes délégations de signature si on fait partie de tel service ou de tel service, on a pas les mêmes responsabilités. Si on reprend par exemple, l'affaire Papon, le respect du fonds est primordial parce qu'en fonction de... enfin lui était à un certain niveau mais au-dessus il y en avait d'autres, etc. Donc ça c'est primordial à comprendre, moi je pense, et à conserver. Là, c'est à nous d'être beaucoup plus exigeants et c'est du coup une des exigences premières, finalement.

Et justement, est-ce que les professionnels, les archivistes, sont prêts à accueillir ce changement et à prendre en compte ces nouvelles exigences ?

Tu as assisté à une réunion, hein ? Donc, euh, oui et non, on va dire. *rires* En fait, c'est qu'il y a une méconnaissance entre... Enfin après je pense que parfois on est un peu trop exigeants avec nous-même, c'est-à-dire qu'on veut des notices d'autorité parfaites alors qu'on est pas là pour faire une thèse sur le service producteur. Pour le coup, il y a des historiens qui sont plus à même de le faire, nous on est là pour amener des éléments de base, au minimum. C'est-à-dire la date d'existence, les textes et règlements etc, on va pas faire une thèse. Donc ça, une fois que ça sera bien rentré en tête... Parce que sinon, les services s'y sont mis, parce que les notices producteurs, on en fait depuis trois ou quatre ans, c'est tout. De façon normalisée, on ne les faisait pas avant, donc c'est un vrai changement, ça, en plus de ce qui se passait. Alors, on a toujours eu une classothèque, qui est très complète, où il y a tous les éléments pour faire les notices producteurs mais sans avoir jamais fait de notices producteur donc là effectivement, c'est beaucoup de reprises du coup, parce qu'il faut les faire. Mais c'est dans le vent aussi, ça va aller dans le bon sens, donc, non, les gens vont s'y mettre et ils s'y mettent. Et puis c'est très drôle, parce que même moi, au début du classement, ce que j'ai fait, quand j'ai commencé à faire du classement sur les fonds de la seconde guerre mondiale, donc c'était les fonds de la préfecture et du cabinet, la première chose que j'ai réalisée c'est l'organigramme de la préfecture, en fonction du temps, pour retrouver qui a fait quoi, etc. Mais finalement, je n'ai pas fait de notices producteur, j'ai juste fait les organigrammes qui nous permettent d'avoir.. Mais si on n'a pas cette vision de ce que fait un service quand on va traiter le fonds, on ne comprend pas les documents qu'on traite, donc l'un des préalables, normalement, de traitement, c'est de comprendre qui est le producteur. Si on ne le fait qu'à la fin... alors, à la fin, on peut le finaliser, mais il faut faire une première trame dès le départ en fait.

Annexe 7c : transcription de l'entretien avec Amélie Antoine

Entretien avec Amélie ANTOINE, archiviste du bureau de la conservation et du classement des Archives départementales de la Gironde. Elle est amenée à travailler de manière ponctuelle sur le projet SAEM, sur les aspects liés à l'archivage des délibérations des assemblées du Conseil départemental.

Entretien réalisé dans son bureau, sur le site des AD rue d'Aviau, le 23/06 en début d'après-midi.

Alors, est-ce que tu peux me parler un peu de toi et de ton parcours professionnel ?

Moi je m'appelle Amélie Antoine, je suis attachée de conservation du patrimoine, cheffe de projet classement et conservation au sein du service collecte et traitement. Je suis rentrée aux Archives départementales de la Gironde en septembre 2011, initialement j'étais plutôt sur la partie collecte, j'étais responsable du bureau des archives de l'État. Entre temps il y a eu une réorganisation et je suis allée, du coup, sur une de mes passions en archivistique, qui est le classement des fonds, pour l'instant beaucoup plus papier qu'électronique mais l'archivage électronique, effectivement, est l'une des préoccupations, aujourd'hui, importantes, des archivistes. Mon parcours après, si on prend le parcours et les expériences professionnelles : des études d'Histoire, un Master Archives, le concours qui est venu un peu plus tard, j'ai d'abord commencé à travailler. J'ai fait des petites missions, comme beaucoup de post-étudiants.

Toujours dans le domaine des archives ?

Toujours dans le domaine des archives, oui tout à fait. J'ai toujours travaillé dans le domaine des archives, dans le classement de fonds, j'ai surtout un profil classement et puis petit à petit je suis rentré dans le domaine particulier de suivi et de sensibilisation des producteurs, aussi bien pour les archives papier qu'électroniques d'ailleurs. C'est très important aujourd'hui, car la définition des archives comporte la notion de tout support, et pas seulement la vision un peu classique qu'on a, où l'archiviste est dans un domaine un peu poussiéreux et un environnement papier essentiellement. Donc effectivement, je pense maintenant avoir un parcours assez complet, les fameux 4C : la collecte, le classement, la conservation et la communication, qui comprend aussi bien la recherche en salle de lecture que les expositions.

Et du coup, ton positionnement ici, aux Archives, c'est quoi exactement, au sein du bureau classement et conservation ?

Alors, je n'ai pas un positionnement hiérarchique, j'ai un positionnement fonctionnel, c'est-à-dire que mon travail c'est vraiment d'être cheffe de projet. C'est-à-dire qu'on me demande... la logique veut que nous ayons une programmation et que selon cette programmation et selon les projets à mener, j'identifie les phases, les moyens et que je constitue des équipes pour mener à bien ces projets.

Et on a eu l'occasion de se croiser à quelques réunions autour du projet SAEM, comment est-ce que tu intervies dans ce projet-là et comment est-ce que tu as été amenée à travailler sur ça ?

En fait, j'interviens parce que, justement, vu que je suis cheffe de projet et que je dois avoir une vision globale de certains projets, et il y en a un qui est très important et intéressant du point de vue archivistique, historique mais aussi politique au sein de notre collectivité qu'est le département de la Gironde. C'est-à-dire, ce que l'on souhaite faire, c'est créer un instrument de recherche, avec une liste en ligne, et un accès du public à toutes les délibérations du Conseil général aujourd'hui, de 1790 à nos jours. Donc forcément, nous nous sommes croisés au cours de réunions, parce que depuis plusieurs années désormais, l'ensemble des délibérations de la collectivité du département sont nativement numériques avec signature électronique. Donc, dans ce cadre-là, et dans le processus d'établissement du profil SEDA propre aux délibérations, aussi bien des séances plénières

que de la commission permanente, nous nous sommes retrouvés autour de la table.

Et sur ces aspects-là, c'est quoi exactement ton rôle ?

Mon rôle, c'est vraiment de définir, et aussi par intérêt parce que je trouve ça passionnant, encore heureux parce que les archives voilà... Mon rôle c'est vraiment de définir, avec la vraie cheffe de projet, c'est Delphine Jamet pour les questions d'archivage électronique, d'établir les besoins. Pour moi, c'est de définir... dans le jargon classique, on dirait le plan de classement. Mais mon but c'est d'identifier les métadonnées à récupérer, comment les récupérer, pour constituer la recherche et mon futur instrument de recherche parce qu'en fait les archives électroniques, fondamentalement, ne changent pas le métier de l'archiviste. Fondamentalement, non, sur des aspects techniques, technologiques, oui. Mes métadonnées, il faut que je les aie, pour décrire, pour faire mon indexation, et il faut que je récupère les documents qui sont créés, donc l'information, les métadonnées, et aussi les documents originaux, c'est-à-dire la question de savoir si j'ai le document définitif, est-ce que c'est bien la délibération signée revenue du contrôle de légalité et revenue de la préfecture, est-ce que j'ai bien la signature, est-ce que j'ai récupéré toutes les pièces, tous les objets, est-ce que j'ai bien récupéré tout cela ? Et ensuite comment structurer toutes ces informations pour je les rassemble dans une notion qui date de 1841, qui est la notion de respect des fonds. Donc moi je suis une archiviste pure, comme l'archiviste du 19e siècle, je n'ai pas changé, par contre je vis avec mon temps.

Et du coup, si tu dis que les outils que les outils sont toujours les mêmes, qu'on se réfère au cadre de classement de 1840, dans quelle mesure est-ce que l'archivage électronique et l'e-administration en général ont impacté les pratiques ?

Oui, ça a impacté les pratiques. Effectivement, je sais que pour beaucoup de collègues, c'est une restructuration très importante, moi pour l'instant je débute un petit peu, je l'avoue. C'est normal, je pense que tous actuellement... il n'y a pas beaucoup de plateformes aujourd'hui qui... nous, il y a la plateforme SAEM qui va bientôt rentrer en production, mais il y a peu encore ça restait théorique. Mais je ne pense pas... Pour moi, fondamentalement, ça ne change pas. Pour moi, il y a toujours des objets, alors que ce soit sur un support électronique... En fait, ce qui change, c'est la question de la conservation. Fondamentalement, ce n'est pas celle de la récupération, c'est plutôt celle de la conservation : comment conserver sur le long terme ces archives, ces données, comment prouver leur authenticité, comment s'assurer de leur authenticité Parce que les archives, il y a deux aspects : il y a un aspect historique, mais il y a un aspect juridique, quand même, il ne faut pas l'oublier. Donc à la limite, s'il y a un changement, c'est comment j'appréhende, comment j'assure l'authenticité, comment je suis sûre que j'ai récupéré toutes les métadonnées, voilà, c'est plutôt ces questions-là qui changent la pratique.

Et justement, tous ces aspects-là, la valeur probante, etc., c'est des choses qu'on retrouve généralement, on va dire, plus dans les pratiques du records management...

Oui, tout à fait.

Est-ce que tu dirais que, ces dernières années, il pourrait y avoir eu une convergence de ces différentes visions de l'archivage, finalement ?

Oui, en fait quand tu travailles principalement dans le records management, que tu es en relation avec des producteurs, l'outil principal c'est le tableau de gestion. Donc, le tableau de gestion, on avait expérimenté ça avec une stagiaire qui est devenue une collègue ensuite, on avait expérimenté ça avec une grande école d'architecture près de Bordeaux et on avait fait un tableau de gestion qui comprenait aussi bien la production papier que la production électronique et c'est là, finalement, que pour les administrations, c'est plus simple d'identifier les difficultés. Alors, je ne parle pas du projet des délibérations, qui là bien sûr sont des documents à conserver, à haute valeur ajoutée, mais souvent dans les administrations, on se fait tout un monde de l'archivage électronique alors que la plupart des données il faut les identifier parce que ce sont aussi des éliminables, il ne faut pas

l'oublier. Par contre, comme pour le papier, je relativise peut-être un peu trop, je schématise peut-être un peu trop, mais comme pour le papier, il faut faire un visa d'élimination aussi, ça reste des données publiques qui doivent passer par le même principe de contrôle technique et scientifique de l'État, pour certaines administrations. Voilà, le système n'est pas révolutionnaire en soi, moi je ne pense pas. Par contre, on a de nouveaux acteurs, de nouveaux collaborateurs, c'est-à-dire des informaticiens, effectivement, comment récupérer le bon format, comment s'assurer que ces informations sont complètes, pérennes, ces questions-là, oui.

Et justement, par rapport à ces nouveaux acteurs, je ne sais pas dans quelle mesure toi tu interagis avec les informaticiens, dans le cadre du projet SAEM ou pas d'ailleurs, mais...

Pour l'instant, pas trop, on a des réunions ensemble mais... Enfin, si, c'est-à-dire que moi je vais chercher à comprendre bien le projet, ce que je souhaiterais apporter, c'est vraiment l'aide pour déterminer le profil SEDA, pour vraiment identifier les métadonnées, c'est plutôt dans ce schéma-là qu'on discute, puisqu'ils m'expliquent ce qui est faisable et ce qui ne l'est pas. Après, la difficulté que je rencontre, c'est le vocabulaire, la difficulté c'est souvent le vocabulaire, c'est-à-dire qu'ils vont utiliser des termes souvent anglophones, voilà : le bus, alors, au début ils parlaient du bus, du bus (rires). Donc voilà, c'est plutôt en termes de terminologie que c'est pas toujours évident pour moi pour le moment, ils se comprennent, ils rebondissent. En tout cas, si je comprends, le temps que je comprenne, j'ai pas le temps d'en placer une, en gros. Mais si je comprends, c'est déjà pas mal. Et j'aimerais beaucoup évoluer vers ça d'ailleurs, surtout sur les profils SEDA, j'ai déjà fait des formations et j'aimerais beaucoup évoluer, c'est quelque chose qui m'amuse beaucoup.

Alors, c'est quelque chose qui t'amuse beaucoup, toi, mais autour de toi, est-ce que c'est une tendance que tu observes ou pas vraiment ?

La plupart des gens n'y entendent rien, ils n'y ont pas été confrontés réellement. Donc je pense que demain, si le référentiel est mis en marche, c'est-à-dire qu'on va nous demander de faire les fiches ISAAR, en EAC, et qu'on utilise le référentiel de manière quotidienne, si on commence à voir... Parce qu'il y a les évolutions du SEDA aussi, il y a des personnes qui sont habituées à l'ancien SEDA, donc les évolutions, s'habituer. Mais là, on a un outil qui nous permet sans difficulté, sans voir « le code », etc., de prendre les choses en main. Je pense que les gens vont s'y habituer, je le pense sincèrement. Après, les interrogations que j'ai, c'est plutôt comment on va faire avec le futur site internet. C'est plutôt sur ces questions-là, parce que moi par exemple sur le projet des délibérations, j'ai un instrument de recherche sur les archives post-1940, j'ai un seul instrument de recherche. Donc d'un côté, on va faire numériser les délibérations, donc là c'est des liens image avec des JPEG dedans, mais comment ça va se passer réellement pour les archives nativement numériques, alors là oui j'ai des vraies questions et je me représente pas encore.

Alors, tu disais tout à l'heure que les outils n'ont pas tellement changé, que finalement on se réfère toujours au même cadre de classement, et justement, ces instruments de recherche, est-ce que pour toi ils sont toujours un outil pertinent, dans un environnement numérique où la donnée, c'est-à-dire la granularité très fine, prend de plus en plus d'importance ?

Tu reviens un peu sur le web sémantique, oui, oui... Pour moi l'instrument de recherche surtout... parce qu'en fait l'instrument de recherche en archives, c'est quelque chose de très complet, surtout si on parle d'instrument de recherche électronique, qui s'en sert totalement dans cet esprit d'échange. C'est-à-dire qu'aujourd'hui, tu as l'instrument électronique de recherche d'un fonds, c'est très complet et tu peux déjà, actuellement, indexer à tout niveau, tu as le niveau haut, qui correspond, si je devais schématiser, à une forme d'introduction, où déjà il y a toute l'introduction ISAD(G) et toute la fiche producteur. Et elle, peut déjà être en lien avec d'autres notices d'autorité, de différents autres organismes, donc finalement je pense que cette pratique... oui, je pense que ça peut totalement s'insérer. L'instrument de recherche, c'est un peu le socle... ce n'est pas comme un livre, l'instrument de recherche donne de l'information dès le début, le plan de classement, que

certain regarderaient et diraient : « oui, c'est une sorte de sommaire », donne de l'information dès le début, et déjà peut être indexé, et déjà il peut être lié à d'autres instruments, par exemple un sous-fonds peut tout à fait être lié à une autre collection d'un autre organisme, qui vient compléter... Tout est inter-relié, parce qu'on a des instruments de recherche électroniques, qui eux, avec un export en XML-EAD et EAC pour les fiches producteurs. Je sais qu'on va avoir RiC-O aussi, prochainement, il y a des évolutions, donc je pense que la réponse est oui.

Et pour les publics... est-ce que ça « fonctionne » pour les publics ? Et justement, tu parlais du site web, comment est-ce que...

Oui, eh bien je ne sais pas trop quoi répondre, parce que comme vous le savez tous, il y a un projet de refonte du site internet et je n'ai pas participé concrètement à ce projet, pour être très honnête. Je sais qu'il y a des projets de recherche fédérée, mais après je n'ai pas été plus associée que cela, je reste un petit peu sur ma faim. Sachant que ça a été repris, il y a peu, par la direction de l'informatique, je n'ai pas plus d'informations que ça.

Parce qu'une des volontés de ce projet de refonte, c'est de faire un différentiel entre les différentes typologies d'utilisateurs et de les prendre en compte directement de l'interface...

Oui.

Est-ce que pour toi, c'est un pas dans la bonne direction, ou est-ce que ça ne risque pas de...

Dans les différents utilisateurs, c'est-à-dire entre le producteur d'archives qui a besoin de... ?

Plutôt le niveau de familiarité avec les pratiques archivistiques.

Ah, là, j'ai raté un peu cet épisode-là moi personnellement. Non, moi je pense que les éléments... Non, moi je ne suis pas pour, pas spécialement. Là, tu m'en parles, je n'ai pas forcément été associée, honnêtement, mais non je ne suis pas pour ? Pourquoi, non, c'est compliqué, c'est alambiqué... Non, ça ne me parle pas, ça. Non, il faut que le système soit ergonomique, effectivement, il faut qu'on puisse faire des tests, c'est pour ça qu'ils veulent le faire selon la méthode agile, faire des tests, mais après non, tu rentres dans une usine à gaz.

Tu avais l'air de dire que des principes comme le respect des fonds étaient toujours clairement d'actualité...

Oui, pour moi, au niveau philosophique, il ne faut pas perdre ça. Par rapport à d'autres métiers comme la documentation, etc. On ne fait pas le même métier.

Et tous ces nouveaux modèles, ces nouveaux standards, par exemple le modèle OAIS pour l'archivage électronique, est-ce que c'est des choses qui, pour toi, correspondent à ces pratiques existantes ?

Oui, complètement. C'est une transposition. Comme on le disait tout à l'heure, le vocabulaire évolue, certes, il faut se l'approprier et que nos cerveaux... mais oui, pour moi il n'y a aucun problème de corrélation, pour les métadonnées, techniques, quand on parle des différentes métadonnées pour l'organisation de l'information. Après, je ne suis pas dessus tous les jours, soyons clairs, mais oui. Par rapport aux normes en vigueur, bon après il y a tout cet aspect normatif aussi à prendre en compte. Le mieux pour moi, ce serait d'avoir un projet où je pourrais aller jusqu'au bout. Pour l'instant, il y a les délibérations, mais on a un peu de mal à... c'est lourd aussi à mener. Il y a beaucoup d'acteurs, par contre, sur ce type de projets, ça c'est vrai. Avant, on pouvait classer un fonds tout seul, à deux, trois personnes, on pouvait se l'approprier, même un tableau de gestion. Là, c'est des projets qui nécessitent beaucoup de monde autour de la table.

Au niveau des types de professionnels différents ?

Oui, c'est ça, c'était dans ce sens-là.

Je sais pas à quel projet tu as participé en dehors du projet SAEM mais est-ce que tu dirais que l'équilibre des forces en présence, entre les différents types de professionnels, est-ce qu'il est plutôt respecté ? Est-ce qu'on a un certain équilibre ?

Ici, pour la Gironde ?

Oui, enfin de ce que toi tu connais. Est-ce que les archivistes arrivent à faire ressortir leur valeur ajoutée, leurs compétences, dans ces projets-là ?

Oui, de plus en plus, dans le cadre de projets de collecte, de versements d'archives électroniques, je trouve qu'il y a quelques années, on était totalement en dehors, et c'est totalement l'inverse aujourd'hui. D'ailleurs, on le voit même dans les annonces d'emploi. On n'a jamais vu aujourd'hui de demande d'archivistes, bon bien sûr archiviste c'est un terme générique, il y a plus d'une centaine de métiers derrière le terme d'archiviste, ça va du webmaster à l'infographiste, à l'archiviste pur et dur au gars qui fait de la paléo sur la série, il y a plein de métiers. Ce que je veux dire par là, c'est qu'on n'a jamais autant vu finalement... je trouve que l'archiviste, il y a encore très peu de temps et même encore dans un sens, était vu comme un vieux monsieur, dans une cave, aujourd'hui on sort totalement de ça. L'archiviste est devenu un acteur à part entière, et je n'aurais pas parié là-dessus il y a encore cinq ans.

D'accord, donc c'est un phénomène récent ?

Oui, et d'ailleurs on voit depuis, on l'a vu... parce que je suis allée au forum, à Troyes, l'année dernière et c'était étonnant puisque pour l'évolution du métier, on s'est retrouvé avec des attachés de conservation, des archivistes qui sont passés, non pas du côté obscur *rire* mais qui se retrouvent à garder leur poste d'archiviste mais à travailler au sein des directions informatiques.

Sinon, est-ce que tu as suivies des formations dédiées à...

Oui, jusqu'à la dernière, d'ailleurs, j'étais déçue, puisque je faisais beaucoup de formation, depuis 2010 je fais quasiment chaque année des formations proposées par les archives de France sur l'archivage électronique. Et il faut être très honnête... à chaque fois je ne suis pas la seule à montrer mon mécontentement parce que jusqu'à présent, on parlait beaucoup des normes, de lois, c'est très juridique tout ça mais ça manquait de concret et on se retrouvait souvent, avec son papier et son crayon, à suivre des Powerpoint. Sauf l'année dernière, où là j'ai fait... où d'abord, à Troyes, je suis allée pour les trois journées consacrées aux données électroniques, où j'avais suivie « comment traiter un vrac numérique » avec Baptiste Nichèle, donc là j'ai trouvé ça passionnant, vraiment. Parce que là ça veut dire que si tu es un peu autodidacte, moi je suis pas très forte mais si tu es un peu autodidacte, tous les logiciels tu peux les télécharger toi-même, tu peux déjà faire un premier travail. Ça c'est ce que j'aime, c'est l'esprit qu'il y a derrière ça, aussi, c'est ça que j'aime bien, c'est-à-dire que l'esprit derrière, c'est quand même le partage d'information, il y a aussi les questions de l'open data aujourd'hui, donc j'ai trouvé ça passionnant. Ensuite, j'ai fait la formation SEDA, en septembre de l'année dernière, toujours avec Baptiste Nichèle. Bon, outre le fait qu'il est très pédagogue, là pour le coup c'était très intéressant et c'était concret. C'était concret, on avait les outils, alors ça peut être un outil libre, mais au moins on a les outils qu'on crée et on montre concrètement... parce que moi je dis toujours, même le papier, c'est pas toujours très simple non plus. Parce qu'on met derrière l'archivage électronique la notion de technique et de difficulté, parce que tu es obligé de travailler avec un informaticien ou du moins une personne dont c'est la pratique quotidienne. Et puis, il sait où sont les documents, parce que souvent tu arrives dans une direction telle que le Département, moi je dis : « mais ils sont où les documents ? », à un moment donné il faut être concret. Et les archives c'est concret, c'est pas du tout théorique, c'est pas du tout philosophique, à un moment donné on a un fonds papier, je veux dire peu importe le support. Et les archivistes, normalement, doivent pouvoir se mettre à la page des évolutions, parce que depuis toujours, l'archiviste est celui qui s'adapte à son époque.

rires

Bon, après c'est vrai que je ne peux pas aller trop loin dans des données plus précises, c'est vrai que Delphine ira beaucoup plus loin, pour le SAEM... Moi je découvre un petit peu, mais j'aimerais bien aller vers ce métier-là. Alors, je ne sais pas si je pourrais faire que de l'archivage électronique, mais ça ne me pose pas de problème, ça ne me déplairait pas. Par contre, il faut que je puisse le faire

quotidiennement, et pas juste une fois tous les six mois. Après, les profils SEDA, c'est super intéressant.

Et pour revenir sur le projet SAEM, justement, comment est-ce que toi tu envisages la place et le rôle du module référentiel ?

Alors, c'est une place essentielle, et c'est là que c'est une vraie révolution, le référentiel. Le référentiel, en fait, permet... Le but, c'est d'avoir un outil qui évite l'erreur humaine, qui évite ce qu'on voit beaucoup dans les services, dès que tu veux indexer, c'est quasiment impossible si c'est réécrit et réécrit et réécrit avec des fautes de frappe, etc. Donc le but c'est d'avoir, aussi bien pour l'indexation, d'avoir comme ça, une sorte de bibliothèque, finalement, et d'avoir aussi bien les termes d'indexation normés, d'avoir les notices d'autorité, très importantes. Parce que là, en plus, il y a une vraie relation, aujourd'hui, au niveau des notices producteur, on est vraiment dans le cœur de notre métier, et le référentiel permet, entre autre, de fédérer tout ce travail, et de le rendre productif, et comment dire... idoine. C'est-à-dire qu'une fois que c'est validé, on n'y touche pas et on peut du coup travailler et aller chercher ainsi les différentes données, comme une sorte de... D'ailleurs, c'est très amusant de se dire, comme ça, qu'on peut alimenter, pour revenir aux profils SEDA, on peut aller chercher ces paquets et organiser ensuite ces informations en évitant, justement, le problème de resaisie, etc. Voilà comment je l'ai compris, moi.

Et pour toi, quel est l'enjeu de faire reposer tout ça, notamment pour certains aspects comme les vocabulaires, sur les technologies du web sémantique ? Comment est-ce que tu le perçois ?

Euh... je ne le perçois pas vraiment comme ça, parce que c'est vrai que le web sémantique, on parle tout le temps de web sémantique, c'est pas toujours évident pour moi de comprendre tout ce que ça recouvre derrière. Parce qu'aujourd'hui on met web sémantique... enfin on dit web sémantique... Non, peut-être que je ne l'envisage pas toujours comme ceci.

Et juste comme ça, est-ce que tu avais pris connaissance de la feuille de route du ministère de la Culture concernant la transition web 3.0, ça te dit quelque chose ou pas ?

Oui, alors je lis beaucoup de documents, je fais quand même de la veille, c'était quand ça...

En 2014.

Oui, j'ai du le lire, encore qu'après j'étais beaucoup sur les services de l'État, parce que là il y a beaucoup de retard aussi sur ces questions-là. Il y a des différences réelles entre les services de l'État et les collectivités, dans l'évolution des programmes. Tu veux qu'on en parle ?

Non, pas forcément, ça aurait été un bon point de départ, mais ce n'est pas grave *rires*

Dis-moi par rapport... si tu m'en parles un petit peu... c'est sur quel sujet que tu voudrais que j'embraye ?

Est-ce que, bon tu n'as pas forcément creusé la question mais est-ce que tu as entendu parler de projets, de ressources, de gisements de données qui ont pris en forme, dans le domaine de la culture, par exemple ?

Plutôt sur l'open data.

Sur l'open data ?

Ouais, au niveau de la direction de la culture, au niveau de la direction de la modernisation, on est plutôt sur des questions d'open data, en fait. Là, pour le coup, ce ne sont pas forcément des sujets qui sont évidents pour moi, là c'est une vraie révolution. Autant, l'archivage électronique, pour moi finalement c'est plus de la pratique et de nouveaux acteurs, autant l'open data c'est une vraie révolution, pour moi en tant qu'archiviste, parce que c'est une autre façon d'anticiper et de penser. Donc, je pense que ça serait plutôt sur des questions d'open data, la mise à disposition des sources, des ressources, des données, c'est plutôt ces questions sur la réutilisation des données, où j'ai plutôt vu la mise en œuvre. Sauf erreur de ma part, c'est plutôt sur ça que j'ai pu constater...

D'accord, et pour revenir sur le nouveau site, comment est-ce que tu vois son... quel est son but, pour toi ? Ou globalement, un site d'archives aujourd'hui ?

Oui, outre la question que tu m'as posée tout à l'heure, ce n'est pas du tout de s'adapter aux différents niveaux mais par contre de différencier, et c'est pour ça que je t'en ai parlé au début, les différents acteurs que l'on a dans la chaîne archivistique. On est d'accord, on va de la collecte jusqu'à la valorisation, donc l'idée c'était de faciliter la communication, puisque aujourd'hui j'ai touché aussi bien aux missions de collecte que de classement, conservation jusqu'à la comm', donc la question c'est plutôt de permettre, alors c'est une vision un peu classique aussi, de permettre l'accès à l'information beaucoup plus important, mais des informations utiles, pas seulement, enfin... D'un côté, tu as la recherche avec les instruments de recherche, ça c'est une chose, mais je suis un service lambda, je suis les impôts, je suis, voilà, je veux pouvoir accéder aux tableaux de gestion, je veux pouvoir accéder à l'information juridique, je veux pouvoir retirer un bordereau de versement ou un visa d'élimination en modèle, je veux pouvoir le transmettre, pourquoi pas avec un code d'accès, via le site internet. C'est-à-dire qu'il faut que le site internet soit une vraie plateforme de travail, aussi bien pour nous que pour les administrations, services versants et services producteurs, pour qu'il y ait un véritable échange à travers ça. Donc ça c'est les choses un peu basiques, mais nous ne l'avons pas au sein des Archives départementales de la Gironde et je trouve que c'est très dommageable. C'est une des choses, pour moi, essentielles, qu'il y ait déjà cet aspect-là, que proposent beaucoup de sites de services d'archives en France. Oui, là ce qui me vient c'est principalement ça, c'est-à-dire qu'on puisse avoir un outil de travail, un vrai outil de travail, c'est-à-dire que ce n'est pas juste un site internet qui donne de l'information, c'est aussi permettre... On parle du projet SAEM, je ne sais pas comment exactement ça va se développer mais l'idée c'est d'avoir une plateforme d'archivage électronique définitive derrière. Donc, il faudrait avoir une partie site tout public et des parties où tu entres un mot de passe, etc. mais permettre aux administrations aussi de travailler avec cette plateforme-là. Mais encore une fois, je le répète, je n'ai pas vraiment été associée...

Non, mais ce n'était pas forcément par rapport à ce projet-là, mais plutôt aujourd'hui, qu'est-ce que ça doit permettre de faire...

D'accord, eh bien je trouve ça intéressant que le site soit un outil, véritablement une plateforme.

Pas un simple espace de consultation, en fait ?

Exactement.

Sinon, tout à l'heure tu évoquais RiC-O, Records in Context, comment est-ce que tu vois son éventuelle influence ?

Pour l'instant, j'ai pris connaissance de ce que ça recouvrait, maintenant pour l'instant, ça reste plus de l'information. J'attends de voir un petit peu, j'attends de voir les évolutions concrètes. Je ne peux pas vraiment en parler.

Donc, tu ne serais pas en mesure de me dire si ça coïncide avec la description archivistique à plusieurs niveaux et la hiérarchie...

Ah, si, normalement si, déjà avec ISAD(G) c'était le cas, normalement avec ISAD(G) c'était déjà le cas, c'était déjà de la description à plusieurs niveaux. Il faudrait que je travaille un peu plus sur cette question, je suis désolée... Moi, ce que j'attends de voir avec RiC-O, c'est une évolution plutôt sur les questions de fonctions, de producteurs, ces questions-là.

Et justement, qu'est-ce que tu attends à ce niveau-là ?

Bon, alors moi je suis une... L'ISDF n'a jamais réellement été mis en place, il n'y a jamais eu de DTD, enfin pas de DTD, mais de langage dédié. Moi ce que j'attends c'est vraiment une évolution au niveau des descriptions, une harmonisation et une optimisation. C'est-à-dire qu'on aille vraiment plus loin, aujourd'hui je trouve que c'est compliqué, on a ISAD(G) d'un côté, on a ISAAR qui fait peur à plupart des archivistes qui ne sont pas des fous de la normalisation. Je pense qu'il faut une harmonisation entre les normes et un outil, enfin une norme, qui reprend l'ensemble des champs,

notamment les champs obligatoires, des différentes normes, d'harmoniser leurs champs...

Tu me prends un peu au dépourvu, je n'ai pas relu RiC-O. *rires*

Non, non, c'est clairement l'un des objectifs du modèle.

Et sinon, il y avait un élément intéressant qui avait été évoqué, je ne sais pas si tu t'en rappelles, au moment de notre formation sur AAC, avec Myriam Pauillac et Louis Colombani, sur le fait, pour revenir sur les sites d'archives, que peut-être les sites d'archives seraient amenés à disparaître et que les données des archives pourraient être amenées à être exposées dans des grands gisements, des portails...

Quand ils parlaient de Google ?

Non, pas forcément, mais oui, ça pourrait être récupéré par des acteurs privés. Et je me rappelle d'une intervention de ta part, tu disais qu'au niveau juridique, ça pourrait avoir des répercussions assez importantes, par rapport à l'environnement qui existe aujourd'hui.

J'ai dit ça, moi ? Si, si, je me souviens, effectivement c'était Myriam Pauillac qui disait que « de toute façon, si vous n'êtes pas vigilants... », moi je me souviens qu'elle avait surtout parlé de Google, où elle disait qu'ils s'attaquaient aux bibliothèques mais qu'ils pourraient très bien s'attaquer aux archives. Après, par rapport à l'environnement juridique, j'étais de bien me souvenir de tous les éléments, moi ce que je voulais dire, c'est que... Archiviste c'est un métier complexe, on doit prendre compte l'aspect normatif, on doit prendre en compte l'Histoire, c'est-à-dire quand on parle d'un plan de classement, c'est pas anodin, c'est vraiment l'articulation... ça parle de l'histoire organique, il y a l'indexation, aujourd'hui on s'ouvre sur le monde et aujourd'hui il y a tout un contexte juridique de communicabilité, de confidentialité et d'anonymisation. Aujourd'hui, l'archiviste, c'est celui qui permet la transparence, ça a une vraie réalité politique, quand on voit le régime des archives d'un pays, on sait à quel régime politique, tout court, on a affaire. On s'adapte mal avec un contexte... c'est-à-dire que l'archiviste doit aussi s'adapter à des questions telles que l'open data, et il faut totalement associer l'archiviste, on ne peut pas s'en passer. Mais je ne suis pas si inquiète que ça, par rapport à ce que pouvait dire Myriam Pauillac, parce qu'à un moment donné la complexité, c'est ça, c'est-à-dire qu'il faut aussi protéger et qui est à même d'analyser un document, d'identifier les textes qui sont liés, ce qui fait que tel document n'est pas communicable, que tel document est à conserver, etc. C'est le travail de l'archiviste, chacun son travail, on n'est pas dans de la lecture publique, les archives ce n'est pas de la lecture publique, c'est tout. Après, ça sert aux sources de l'Histoire, mais on n'est pas dans de la lecture publique. C'était surtout dans ce sens-là que je voulais le dire. Oui, voilà, il faut quand même rester vigilant, c'est important.

Sinon, j'avais une dernière question sur un autre aspect, je ne sais pas si tu as entendu de parler de ça, si tu es familière avec ça : le concept d'Humanités numériques, d'Humanités digitales...

Non, pas du tout, jamais entendu parler de ça. *rires*

D'accord, et si je te demande comment tu vois les enjeux au niveau de la recherche, pour les chercheurs, de la mise en ligne et de l'échange de documents d'archives numériques.

Par rapport à la recherche ?

Oui, comment est-ce que ça a fait évoluer ce domaine-là, finalement ? Est-ce qu'il y a de nouveaux enjeux avec cette diffusion numérique ?

Pour le coup, c'est un sujet que ne je connais pas, donc c'est bien, je pourrai rechercher les évolutions. Eh bien, moi en tant qu'archiviste, j'ai plutôt vu les évolutions pour la collecte des archives des chercheurs, il y a toute la question de la récupération de leurs données, ce qui est relativement récent, parce qu'il y a toujours la difficulté de savoir si les fonds sont publics ou privés, et ils ne sont pas toujours prompts à transmettre leurs documents, donc ça c'est des questions en termes de collecte et de difficultés de la récupération des données. Ça me fait plutôt penser à ça de prime abord, après quels sont les enjeux que recouvre cette notion d'échange... Moi, c'est plus en termes de collecte qu'on a beaucoup de difficultés, en tant qu'archivistes, mais ça évolue bien.

C'est-à-dire que là, sur les données de la recherche, on arrive à sortir d'un no man's land. Mais ces projets-là, les archives des universités, non ça ne me parle pas, je vais aller regarder. Non, en termes de collecte, on a de grosses difficultés à collecter les bases de données, etc., parce que de toute façon il y a une vraie réflexion sur tout ce qui est archives publiques, archives privées, archives de la recherche, des universités, là aussi c'est en train d'évoluer. Et là aussi, il y a un rôle des Archives, en termes de relais, et des archivistes. De toute façon, nous aussi finalement, d'une manière ou d'une autre... comme je le dis souvent l'archiviste c'est un chercheur, mais des institutions, c'est un spécialiste des institutions.

Je crois que j'ai fait le tour de mes questions...

Après, comme je t'expliquais, moi j'aborde certaines questions, aujourd'hui, de l'évolution. On dit l'évolution du métier, mais pour moi c'est une évolution... il y a des adaptations, en termes de compétences à acquérir, mais fondamentalement ça ne change pas mon métier. Après, je ne suis pas une spécialiste, je ne peux pas rentrer non plus dans certains domaines du web sémantique mais...

Oui, mais je veux dire... je t'ai déjà un peu posé la question, mais toi tu fais preuve d'une certaine curiosité vis-à-vis de tout ça, mais est-ce que tu ne crois pas que, finalement, la culture professionnelle, ou je ne sais pas, des archivistes, voir même les formations, ont un rôle à jouer sur...

Oui, bien sûr. Oui, complètement, mais pour moi l'évolution est déjà en marche. L'évolution positive... je dirais qu'il a fallu dix ans, mais il y a dix ans on n'était pas prêts, personne n'était prêt. Tout le monde en parlait, c'est pour ça que j'ai suivie des formations, je crois que ma première formation sur l'archivage électronique ça date un peu, de 2010 ou 2009, ça n'apporterait rien en soi. Mais l'adaptation du métier aux évolutions... il a fallu des exemples concrets, il a fallu la mise en œuvre concrète, aujourd'hui on parle de carte mentale, il y a quand même des concepts qui arrivent et dont on ne parle que depuis l'année dernière, j'ai envie de dire. Après, c'est vrai qu'il y a Delphine Jamet qui est vraiment engagée sur toutes ces questions-là, qui s'est spécialisée et qui nous amène certains aspects « prédigés », parce que quand tu vois le référentiel, pour un archiviste qui n'est pas du tout formé, il faut vraiment qu'il mette de la mauvaise volonté pour ne pas comprendre certains aspects, quand même. Parce que c'est comme une aide, une aide incroyable.

Pour toi, c'est un outil qui fait sens, qui ne pose pas de difficultés... ?

Oui, après je ne l'ai pas pratiqué en long et en large, j'ai déjà fait un profil SEDA, j'ai déjà fait une notice d'autorité. J'ai regardé les différentes fonctions mais pour l'instant ça reste jeune, oui. Ça reste quand même un projet... ce n'est pas encore en production, il y a eu pas mal d'évolutions récemment aussi, il y a eu des bugs, c'est normal, le projet vit. Ce n'est pas une critique, loin de là. Après, est-ce que pour toi, le fait que tu insistes beaucoup sur le référentiel, est-ce que c'est parce que c'est une façon d'appréhender et un projet qui sort de l'ordinaire, parce que les autres projets n'ont pas développé ce type d'outils. On parlait des formations, moi j'ai suivie d'autres formations, tu sais il y a des retours d'expérience de ce qui est fait par les autres régions, et ce type d'outils, c'est novateur. Ce type d'outil a pris en compte le fait qu'il faut une sorte de plateforme qui permette de réunir toutes les ressources, de les fédérer, et d'éviter le caractère qu'on voit dans tous les services, qui est l'erreur humaine et de se retrouver avec des listes d'autorités qui ne sont pas propres, qui ne sont pas normées ou finalisées. Ça permet surtout d'avoir un outil pérenne, surtout. Donc pour moi, je vois totalement l'intérêt. Et justement, pour des archivistes qui dont du mal, pour qui ce n'est pas – j'allais dire leur génération, pardon – je pense que ça peut être un outil qui va permettre justement de passer le pas de l'appréhension. Il y a ça aussi, ça c'est important, je l'ai vu comme ça, c'est un outil qui permet aussi d'être un relais, d'être un outil d'appréhension, de compréhension et de préhension. Pas d'appréhension, mais de préhension, c'est-à-dire que tu t'appropries des procédures, parce que ça, ça rejoint notre pratique quotidienne, sur d'autres outils. J'avais fait aussi des formations qui étaient fort intéressantes, mais ce n'était pas sur l'archivage

électronique, tel qu'on l'entend, avec le développement des plateformes, etc., c'était une question que je me posais et j'aimerais bien en discuter plus souvent avec d'autres personnes, d'ailleurs, c'était de dire : « demain, l'archivage électronique, si tu n'as pas de plateforme ou si tu n'as pas fait un bordereau normalisé type SEDA, imaginons... » Enfin je veux dire, s'il y a un vrai débat à mener aujourd'hui, c'est que l'archiviste n'archive pas, ne fait pas d'archivage électronique et n'en parle pas parce qu'il ne peut pas. Bah non, si, tu peux, tu peux en parler, tu peux organiser tes données, on parle de records management mais dès le début tu peux organiser tes données, tu peux faire en sorte que tes données électroniques vitales soient identifiées, voilà, le travail électronique existe dès l'amont. Éviter ce qu'on appelle le vrac électronique, avoir des métadonnées, des procédures.

Et justement, tout à l'heure, je te parlais de convergence des approches entre records management...

Oui, totalement, c'est là que tout commence.

Mais finalement, est-ce que ce n'est pas quelque chose qui est engagé depuis longtemps, en termes de pratiques, pas forcément en termes d'environnement, est-ce que ce n'est pas quelque chose qui, en France, est engagé depuis longtemps ?

Oui et non, dans la pratique professionnelle, oui. Dans les administrations, c'est autre chose. Après, je ne peux pas parler pour toutes les administrations mais aujourd'hui on conduit dans les administrations... déjà même parfois l'archivage papier c'est [compliqué], il y a encore du travail à faire sur comprendre la notion de producteur, comment j'organise les archives, pourquoi j'archive annuellement, etc. Donc, c'est comme les maths, il faut quand même la base. Mais après, disons que, l'informaticien, finalement, même s'il n'entend pas l'information comme nous, il a déjà ces pratiques-là, il comprend ces informations-là. C'est pour ça que maintenant c'est le couple infernal. Mais l'un ne va pas sans l'autre par contre. Parce que parfois j'ai eu des messages d'informaticiens qui me disent « ça fait cinq ans, c'est plein, on vide tout ». Bah non, non. Donc effectivement, à un moment donné, il faut savoir conserver... C'est ce que je disais tout à l'heure pour l'archivage électronique, on oublie très souvent... on va stocker, c'est-à-dire que l'informaticien va stocker, je suis désolée, je reviens sur la conservation à long terme, mais il va stocker. Il parle de poids, il stocke...

Il voit ces choses-là en termes de poids, de sécurité, de...

C'est tout. Bah, c'est tout. S'il veut bien faire, à la limite il va stocker tout ce qu'il peut, il va stocker, il va stocker. Mais on ne lui demande pas toujours de stocker. C'est-à-dire qu'il y a des informations numériques qui peuvent disparaître, réglementairement, ça peut. J'avais suivi une formation, justement, c'était intéressant, c'est comment mener un audit pour mettre en place un système d'archivage, mais plutôt de type records management, dans un grand ministère, je crois que c'était pour le ministère des Affaires étrangères, et en fait la fille a commencé par identifier tous les grands logiciels existants, tous les types de bases de données existants, tout ce qui était produit numériquement pour arriver à faire une sorte de tableau de gestion, en gros, et aussi pour l'arriéré. Parce que tu as beaucoup... que ça soit du papier ou de l'électronique, ça reste de l'arriéré, ou vrac dans certains cas. Et de dire à un moment donné : « la base de données est ancienne, vous n'allez pas pouvoir récupérer ça en format pérenne mais en même temps ça peut être de l'éliminable », « Ah bon c'est possible ? ». C'est-à-dire qu'il faut quand même analyser l'information que tu conserves. Par contre, éviter de créer de nouvelles bases de données, ou d'acquérir des logiciels, qui ne permettent pas l'extraction, en vue de la conservation de l'archivage, etc. Donc en fait, aujourd'hui l'archiviste doit être, mais ça c'est déjà une pratique connue et reconnue, l'archiviste doit être présent dès la création, quasiment, du document, si je devais résumer. Dès sa création, il faut organiser son information, les métadonnées, le nommage, et savoir identifier à terme ce que va devenir ce document. Parce que ça coûte cher aussi, ça coûte même très cher, de conserver et de stocker des données.

Et tout à l'heure, je te demandais si les archivistes faisaient des efforts pour comprendre les notions techniques, informatiques et tout ça, mais dans l'autre sens, est-ce que toi, à travers le projet SAEM et plus globalement, tu observes des efforts de la part des informaticiens ?

Oui, du peu que j'en ai vu, oui. En tout cas, les personnes qui sont associées à ce projet, mais après tu sais j'ai été là de manière sporadique, mais les personnes que j'ai vues associées à ce projet... D'ailleurs, j'ai été très agréablement surprise, parce que ça ne se passe pas partout comme ça, soyons clair. Après, il y a Pascal Romain, aussi, qui est quelqu'un qui a quand même travaillé aux Archives, donc c'est lui le lien pour beaucoup de choses.

Donc pour toi, c'est important cette double casquette ?

Et bien effectivement, est-ce qu'il n'y est pas pour beaucoup ? Pas dans le projet... après, chaque acteur apporte sa pierre à l'édifice, mais finalement on a quand même quelqu'un qui comprend toutes les notions que l'archiviste lambda ne comprend pas, qui est un technicien, mais qui en même temps a travaillé aux Archives départementales, comprend notre logique de travail, comprend le monde des Archives et lui... c'est une sorte d'hybride, lui, en quelque sorte, est-ce qu'il n'y est pas pour beaucoup ? Il s'était présenté, aux journées de l'AAF à Troyes, parmi ces archivistes qui ont évolué dans leur métier et les nouveaux métiers de l'archiviste, ce n'est pas anodin.

Et puis dans le COPROJ, il y a cette organisation en binômes, aussi, qui y est peut-être pour beaucoup.

Oui, oui. Après, ce qui était riche aussi, c'est que c'est pas seulement le Département, il y a la métropole, j'ai trouvé ça très riche et intéressant d'avoir tous ces acteurs autour de la table. D'autres collectivités aussi, puisque du coup on se pose des problèmes avec des fonctionnements différents. Je voyais l'autre jour, on s'est pris la tête pour arriver à comprendre comment l'information, comment le document, par quelles étapes il passait pour le contrôle de légalité, ce que la préfecture envoyait, est-ce que c'était un mail, est-ce que c'était un mail ou un dossier. Est-ce que c'était un fichier et est-ce qu'il était à conserver ou est-ce que c'était un éliminable ? C'est des vraies questions. Et du coup, il y avait un gars de la métropole qui y avait réfléchi, nous on y avait pas forcément réfléchi, et du coup ça fait un échange d'information qui est intéressant. Et après on a aussi la relation purement financière, commerciale aussi, avec les boîtes privées. Il y a ça aussi, le financement, parce que quand même derrière tout ça, il y a un coût. Derrière, il y a le projet AD ESSOR pour les subventions.

Et justement, tu parles, je suppose, des prestataires et éditeurs de logiciels, là aussi il faut peut-être que les archivistes soient en mesure de dialoguer directement avec eux pour exprimer leur...

Oui, pour exprimer leurs besoins, oui, je suis d'accord.

Est-ce que c'est quelque chose que toi, tu...

Je ne l'ai pas souvent fait, surtout sur ces questions-là. Mais est-ce que ça me ferait peur ? Non, je ne pense pas. Bon après, quand tu es client, eux aussi vont des efforts pour s'adapter, je pense que ça doit être faisable. Moi, sur tous ces projets, que je trouve fort intéressants, j'y suis de manière ponctuelle, donc je ne pourrai pas te dire exactement toutes les difficultés. Mais je les ai vus en réunion, pour le SAEM, ils ont quand même réussi à avoir une culture commune en termes de communication, tous. L'informaticien comprend certains mots de l'archiviste, l'archiviste... on voit qu'on puise dans une seule méthode, la méthode agile, voit que... Il aurait fallu voir au début, peut-être les deux premières séances, je ne sais pas. *rires* Mais là, maintenant, on voit qu'ils sont rodés, quand même. Après, comme je t'expliquais tout à l'heure, je suis très honnête, moi je comprends ce qu'ils disent, tous, mais le temps que ça monte au cerveau, honnêtement, ce n'est pas moi qui vais dire « ah mais oui, moi je pense que... », non. Après, les prestataires, effectivement, il y a une logique commerciale très forte et financière, en fait tu passes la moitié des réunions sur ces questions-là. Là, je sais qu'ils veulent développer un connecteur pour l'export de AIRS Délib, qui est le logiciel qui permet d'éditer, etc., *souple* mais ce n'est pas inintéressant, pour le coup c'est

une découverte pour moi aussi de voir comment tu peux dialoguer. En même temps, là c'est un cas particulier parce que Digitech ont un quasi-monopole sur ces questions-là. C'est pas évident du coup, mais c'est intéressant et la dernière fois c'était plutôt houleux, on va dire.

Et ça fait plusieurs fois que tu mentionnes les méthodes agiles...

Oui, parce qu'en fait c'est une méthode qui est vraiment propre au monde des informaticiens, j'en avais déjà entendu parler auparavant, mais ce n'était pas quelque chose qui était... ce n'était pas quelque chose que j'entendais dans le monde des archives, et même le mode projet. C'est-à-dire qu'aujourd'hui, moi je suis cheffe de projet, on travaille en mode projet. On a des manières, des nouvelles organisations du travail qu'on ne connaissait pas avant dans le monde des archives. Donc là en l'occurrence il y a une imprégnation, une acculturation managériale, au niveau des procédures, des « process ». Mais c'est normal, justement tu parlais d'évolution des métiers, c'est-à-dire qu'à un moment donné, il y a un dialogue entre les différentes institutions, que ce soit au niveau normatif, juridique, technique. Je pense que les archivistes ont un peu envie d'aller sur le terrain des informaticiens, je pense que les informaticiens ont un peu envie de comprendre tout l'impact juridique, parce qu'au sein de leur collectivité, c'est une valeur ajoutée aussi. Donc tout le monde arrive à se parler et la méthode agile aussi... on se rend compte que c'est des projets qui ne sont quand même pas évidents et que finalement la souplesse permet d'organiser des séances de travail, c'est intéressant. Après, je ne suis pas une spécialiste non plus.

Eh bien écoute, merci beaucoup.

Avec plaisir, je suis désolée... Après c'est vrai que sur le web sémantique, j'ai un souci avec le web sémantique. Je comprends ce qu'est le web sémantique mais en même temps je trouve qu'on met le web sémantique à toutes les sauces. Quand on te parle de web sémantique, au-delà de Wikipédia... Pour moi, la mise en relation des informations... C'est vrai qu'aujourd'hui, on le voit avec les gamins, moi je suis vieille j'ai 40 ans, toi t'es jeune... Mais les gamins, ils ne cherchent pas comme moi je cherche. Ça ne veut pas dire que je ne comprends pas la logique, mais il y a une nouvelle façon d'accéder à la connaissance, c'est étonnant. C'est étonnant ce système de relations entre les données, c'est super. Tu parlais de la recherche, mais aujourd'hui, la recherche de manière globale, l'accès à la connaissance, c'est incroyable. C'est passionnant de travailler sur des métiers comme ceux-là, nos métiers. La rapidité, aussi, à laquelle tu accèdes...

Donc, tu vois bien les enjeux...

Oui, bien sûr.

Mais est-ce que tu arrives à mettre en relation cette nouvelle vision du web avec les archivistes et leurs...

Oui, complètement, mais ça c'est quelque chose qu'on a à l'esprit depuis longtemps, même sans parler de normalisation. Même quand tu normalises un descripteur géographique, tu fais en sorte d'identifier... de te dire qu'aujourd'hui, c'est une plateforme internationale, qu'il faut accéder à l'information, qu'il faut pouvoir rebondir, comment dire... Tu vas mettre Paris, il y a beaucoup de Paris dans le monde. Voilà, ce sont des données qui ne sont pas propres au métier d'archiviste, d'ailleurs, soyons clairs. Parce que les archivistes, sur ces questions-là... si on demande aux bibliothèques / documentation, il ne faut pas l'oublier. Par contre, nous, notre particularité, c'est quand même cette notion qui est que l'archive, c'est unique, que l'archive a une histoire organique et cette notion de respect des fonds, tu ne peux pas poser les questions de la même manière que pour les autres métiers : bibliothèques, documentation, informaticiens, bon l'informaticien c'est différent, il évolue dans une structure... Il y a ça aussi, il ne faut pas oublier que l'archiviste, même s'il travaille avec l'informaticien, à un instant T... Enfin, moi, mon métier c'est quand même de permettre la sauvegarde de la mémoire et l'accès aux sources. Le but ultime, c'est ça. Je suis très honnête, on fait notre travail, on ne peut pas tout conserver, on donne notre analyse, nos compétences juridiques pour analyser ce qui peut être éliminable, mais quand même, à terme, moi

ce que je veux c'est préserver et sauver des documents de l'oubli, de permettre l'accès, de permettre la conservation de la mémoire. Tu n'es pas archiviste pour rien, l'archiviste a un but, de permettre l'accès à l'Histoire, après je ne vais pas rentrer dans tout ce qui est historiographie...

Pour toi, ce qui est important, c'est l'accès à ces sources-là, mais est-ce que c'est l'accès en tant que documents ou les informations qu'on y retrouve ?

Non, les informations... tout, tout ce qu'on y retrouve. L'archiviste, ce qui est important pour lui, c'est organique, c'est la structure, l'histoire du producteur, ses évolutions, un peu comme organigramme. Tout ça c'est important, tout ça, ça parle, ça a du sens. Et il y aussi le contenu du document, bien sûr, aussi. Mais c'est vrai que quand tu penses papier, tu penses tout de suite... tu identifies directement, tu vas lire les informations dessus directement, tandis qu'un document numérique, la matérialité n'est pas la même, ne serait-ce que cela. Non, la donnée en elle-même et les métadonnées ont du sens. Par contre, il faut arriver à les percevoir, parce que clairement, entre le papier et l'électronique, le sens du métier est le même. Il n'y a pas de différence en soi, le but est le même, c'est ça que j'essaie de dire. Après, il faut s'adapter, il faut que tu arrives à rechercher... à mener ton travail bien, différemment, mais l'essence est la même. Mon but, c'est que les délibérations soient... Il y a aussi une notion de sauvegarde, l'archiviste il veut sauver, il veut conserver et sauver, il veut permettre l'accès. Mais tu as des archivistes qui sont un peu, comment dire... Il y a une vieille école, pour eux ce qui compte le plus c'est la sauvegarde, la conservation, mais moi non, c'est la communication, la transparence, l'accès, mais voilà, il y a aussi l'idée de sauver les documents, leur intégrité, qui elle, est aussi importante en termes juridiques, accessoirement. *rires*

Annexe 7d : transcription de l'entretien avec Damien Dudezert

Entretien avec Damien DUDEZERT, responsable d'application gestion documentaire et archivage, agent de la DSIN. Pas intégré au projet SAEM mais a participé à la rédaction du cahier des charges pour le nouveau site des Archives départementales et est en contact avec les différents éditeurs de logiciels pour assurer le maintien des outils de gestion métier.

Alors, déjà, est-ce que tu peux commencer par me parler de toi et de ton parcours ?

De mon parcours en termes de...

Ton parcours universitaire, tes études, ton métier...

Alors, initialement, j'ai un BTS, obtenu en 2001. Suite à ça, j'ai travaillé, j'étais analyste programmeur à l'association des paralysés de France puis chef de projet développement à Paris, à l'association des paralysés de France. Et donc, il y a cinq ans, je suis rentré ici comme responsable d'application gestion documentaire. J'ai repris mes études, en cours du soir, pour aller récupérer un bac +3 et un bac +4 et j'ai passé une reconnaissance d'expérience professionnelle pour ingénieur.

Ok, et donc tes fonctions ici, aujourd'hui, ça implique quoi exactement, au quotidien ? Tu travailles sur quels projets ?

Alors, je m'occupe de trois directions : la direction des archives, la direction de la documentation, la bibliothèque départementale de prêt, je travaille pour le service des assemblées également, une partie du courrier et des sujets un peu transverses. Donc, en fait, je m'occupe de la maintenance applicative, évolutive et correctrice et aussi d'autres projets quand il a de gros soucis.

Du coup, ça t'amène à travailler avec qui, comme type de structures, tu parlais de la BDP, les archives aussi...

Pour la gestion documentaire les archives, la doc, la BDP dans une moindre mesure, voilà à peu près. Après, c'est pour la démat', pour le service des assemblées c'est les délibérations qui sont dématérialisées, et pour le courrier, ça sera une dématérialisation du courrier réservé.

Ok, et comment est-ce que tu as été amené à travailler sur ces aspects ?

C'est comme ça.

C'est comme ça ? Tu as été nommé sur ce poste, tu n'avais pas de formation spécifique à ces aspects-là ?

Voilà... pas du tout.

Du coup, ces interactions avec ces différents services, ça t'amène à interagir avec quels types de professionnels ? Par exemple avec les archives...

Alors, il va y avoir tous les prestataires de logiciels pour lesquels nous avons des marchés, donc ils ont tous un outil de gestion. Les Archives ont Arkhéïa, la BDP ont Orphée, la Doc ont CADIC. Pareil, pour le service des assemblées, on a Digitech comme prestataire. Donc c'est souvent beaucoup de prestation, de marchés qui vont avec.

Mais tu ne travailles pas directement enfin pas tant que ça, avec les archivistes ?

Je travaille pas par rapport à leurs... enfin...

Plutôt sur leurs outils ?

Oui, sur les outils, mais pas directement sur leur cœur de métier par contre. Là, je vois à peu près ce qu'ils font, j'ai appris à connaître leur quotidien, mais par contre j'ai pas d'influence sur leur quotidien.

Tu n'as pas d'influence mais est-ce que ça veut dire quand même que tu dois maîtriser ou au moins connaître certains concepts, un vocabulaire précis pour pouvoir dialoguer avec ces

personnes-là ?

Alors, le vocabulaire est important. Il faut le comprendre effectivement, mais c'est juste de la compréhension, j'ai pas besoin d'aller plus loin, je ne pourrai jamais être archiviste, je ne pourrai jamais être un documentaliste, voilà.

Et est-ce que toi, par contre, ça te pousse à adopter un, sur les aspects plus techniques, quand tu t'adresses aux archivistes, un discours... enfin à adapter ton discours au niveau technique, ce genre de choses ?

Oui, oui, effectivement, dans les échanges, je ne parle pas de la même manière aux archives, à la documentation ou même à la BDP. Par exemple, la BDP, on est vraiment sur... on ne parle pas du tout aux métiers des bibliothèques, quoi, à part des termes comme notice qui vont revenir, UNIMARC un petit peu, mais c'est vraiment parce qu'on sait de quoi on parle, sinon on est sur un langage tout à fait normal. Après, pour les archives, [il y a la notion] d'EAD, mais sinon pas tant que ça.

Donc là aussi, tu connais les termes, mais...

Ouais, la première couche, quoi. Après, je ne suis pas bilingue EAD par exemple

Donc, de ton point de vue, tu ne peux pas vraiment voir dans quelle mesure les archivistes ont à faire évoluer leurs pratiques avec le développement de l'e-administration et du numérique en général ? Est-ce que c'est quelque chose que tu peux ressentir de ton point de vue ?

Alors, non, je dirais que non, je ne le vois pas vraiment, quoi. En fait, je m'adapte à leurs situations, s'ils ont un besoin et qu'ils me demandent de les accompagner, je les accompagne. Mais après, dans leur quotidien directement, non je ne me rends pas compte de tous les bouleversements liés à ça.

D'accord, et tu ne ressens pas particulièrement de montée en compétence de leur part, ou ce genre de choses ?

Si, avec certains profils. Avec certains profils, on discute de manière un peu plus approfondie sur le sujet. Après, je ne rencontre pas tout le monde aux Archives non plus, j'ai trois ou quatre personnes référentes aux Archives, deux personnes avec qui je vais parler un peu plus, comme Delphine Jamet qui justement sur ces questions du numérique est un peu plus avancée donc on partage un petit peu plus.

Est-ce que tu peux revenir rapidement sur ce que ça implique, quand tu travailles avec ces différents prestataires sur les outils de gestion. Qu'est-ce que ça implique en termes de mise en œuvre, qu'est-ce que tu fais exactement ?

Alors, je fais de la maintenance corrective, donc c'est moi qui sers d'intermédiaire aux agents qui ont des difficultés et aux prestataires. Donc, il faut que je remonte aux prestataires les difficultés qu'ils rencontrent, et après c'est moi qui suis chargé du bon déroulé de la correction, en fait. Je dois paramétrer notre environnement technique pour pouvoir faire correctement les mises à jour.

Et c'est toi qui installes les mises à jour, ce genre de choses ?

Oui, parce qu'on a une politique qui fait que... Déjà c'est bien si moi je peux le faire, voilà, parce que je n'ai pas forcément le droit de le faire, et même les prestataires parfois ne peuvent pas toucher à certains trucs.

Donc, tu m'as dit, aucune formation à propos de l'archivage ou quoi que ce soit ?

Non, rien.

Ok.

Et au niveau du web sémantique, est-ce tu es familier, un peu, de ces notions-là ?

rires Alors, j'ai des notions, parce que je m'y suis intéressé, histoire de comprendre le principe,

mais après *soupir*...

Mais tu t'y es intéressé personnellement, là aussi, rien de plus ?

Oui, je m'y suis intéressé parce que le sujet revenait souvent, par rapport aux archives et à certains échanges. C'est des choses, aussi, que j'ai vu dans ma formation, comme je le disais au début, quand j'ai repris mes études il y a deux-trois ans, où justement on a parlé un peu de web sémantique.

Et du coup, est-ce que ça t'a permis de percevoir l'intérêt que ça peut avoir pour la mise en œuvre de projets de gestion documentaire, pour les archives ?

Oui, oui, je vois bien, après je ne me rends pas forcément compte de la différence que ça peut apporter, enfin je la vois concrètement, mais je ne sais pas comment l'exprimer. Je ne vois pas jusqu'où ça peut nous amener en fait.

Et est-ce que c'est une difficulté, à percevoir ça, que tu as observée autour de toi, dans tes différents échanges, que ce soit avec les archivistes, les bibliothécaires et même les prestataires ?

Alors, au niveau des prestataires, le dialogue avec les prestataires, il est compliqué, parce qu'un prestataire ça sait tout faire, mais si ça ne sait pas le faire, donc c'est très compliqué. Souvent, tant qu'ils n'ont pas un vrai projet à réaliser, ils ne se lanceront jamais, forcément, puisqu'ils ne capitalisent pas dessus, donc ce n'est que si nous on lance quelque chose, que eux montent en compétence aussi. Donc je sais qu'on a un prestataire qui est plus dedans, qui en peut en parler pendant des heures et avancer sur le sujet. Et après, euh... Tu m'avais dit les prestataires, et qui d'autre, dans ta question ?

Les autres catégories de professionnels, si t'as repéré cette difficulté à se rendre compte de ce que ça implique concrètement ?

Oui, oui, tout à fait. Alors, il y en a qui maîtrisent, on sent que quand on parle avec Delphine, ou même Pascal, il y a une maîtrise sur le sujet. Après, il y en a qui perçoivent ce parcours de données, qui paraît plus pertinent, mais voilà, concrètement, tant qu'on n'aura rien de visuel, dans les moteurs de recherche, on ne pourra pas vraiment se rendre compte.

Tu viens de me dire que tu avais du mal à l'exprimer, mais est-ce que tu peux essayer de me dire en quoi ça pourrait être utile, justement, pour les archivistes ou pour les publics ?

Alors, moi je le vois plus pour les publics, pas vraiment côté archivistes, en fait. C'est là où, justement je ne vois pas le cœur, peut-être, après j'imagine que c'est la manière dont ils créent leurs données et dont ils vont mettre à disposition les métadonnées qui est importante. Par rapport aux publics, je vois surtout une ouverture dans la recherche, c'est faciliter la recherche, c'est... une ouverture de la recherche, ils vont lancer une recherche, ils vont pouvoir découvrir plein de sujets qui vont aller à côté, qui vont être connexes et qui vont permettre d'avoir quelque chose de bien ciblé et de construit, en fait. Voilà, c'est comme ça que je le perçois. Par contre d'un point de vue production, c'est là que je ne vois pas trop comment ils font.

Du coup, est-ce que tu as eu des échos, est-ce que tu as connaissance de projets qui reposent sur le web des données ?

De projets... ici ?

Non, de manière générale.

Non, je vois les projets que nous avons en cours ici, après je ne fais pas de la veille spécialement là-dessus. Je sais, si, quand il y a un prestataire, pour le passage de l'EAD aux données pour le web sémantique, mais voilà quoi...

Du coup, tu as travaillé sur le cahier des charges pour le site ? La refonte du site des archives. Absolument.

Et Pascal m'a dit que, initialement au moins, dans la façon dont était pensé ce site, il y avait des éléments orientés web sémantique.

Alors, ce qu'on a fait, comme il n'y a pas de réelle maîtrise dans le groupe projet, ce genre de choses, et qu'on a besoin d'un produit qui sorte assez rapidement, on est parti sur un périmètre initial qui va recouvrir ce qui se fait actuellement, avec l'EAD, ce genre de choses, et on a une ouverture, à l'intérieur, pour le web sémantique. En gros, on a un premier lot, qui va faire un truc assez standard, qu'on possède actuellement, et le prestataire, dans sa réponse, devra nous montrer, nous prouver, qu'il peut nous accompagner, justement, au passage au web sémantique. Mais ce n'est pas la priorité du premier lot, parce que justement quand on a essayé de voir ce que ça donnait et quelques données, on n'avait pas de quoi écrire un cahier des charges spécifiquement, orienté web sémantique.

Pour revenir sur ce premier lot, c'est quoi les nouvelles priorités, les nouvelles attentes par rapport à la première version du site ?

Alors, actuellement, le système des Archives repose sur trois applications. Donc, on a l'application GAEL, qui est l'outil de diffusion des données de l'EAD, on a le site institutionnel et on a la bibliothèque, donc tout ça c'est trois outils différents, trois façons d'accéder à de l'information pour l'utilisateur. Le premier objectif c'est d'avoir un seul outil, c'est-à-dire que l'utilisateur, il veut faire une recherche, il va faire de la recherche dans la partie CMS, dans l'expo virtuelle... On voulait, cet éclatement d'outils, on veut le centraliser, ça c'est l'objectif premier. De plus, ce site institutionnel, il est plus ou moins mort, il est techniquement mort, en fait on ne peut rien faire, donc on se contente de partager des articles, et encore ça ne marche pas à chaque fois. Donc, le premier objectif, c'est d'avoir un site contemporain, avec un design et une accessibilité plus simples, et faire de la recherche unique à travers cet outil... c'est la recherche fédérée. En amenant, quand même, à l'utilisateur, une cinématique plus simple que ce qu'on a fait jusqu'à maintenant, essayer d'arriver à séparer la partie professionnelle, avec tous les gens qui connaissent le langage archivistique, et l'utilisateur standard, qui a envie de faire une recherche sur l'état civil, pour qu'il ne soit pas perdu dans un cadre de classement, qu'il ne soit pas obligé de passer par un instrument de recherche...

Et est-ce qu'il y a une différenciation entre les différents profils, entre les différentes typologies d'utilisateurs, qui est prévue ?

Oui, tout à fait, on a essayé de mettre... Justement, on a fait, oui c'est ça en fait, la partie "pro" et la partie... Professionnel, ce n'est pas au sens de professionnel des archives, mais au sens de personne qui maîtrise le domaine, qui lorsqu'ils tombent sur un cadre classement, *pof pof* ils cliquent, et ça arrive, quoi.

Et est-ce que tu sais comment ça doit se matérialiser, ça ?

On ne sait pas trop, c'est pour ça qu'on part sur une forme agile, pour le cahier des charges, où on dit : « voilà, on voudrait faire ça » et on va le faire en itération, histoire de trouver la meilleure solution, en fait.

Donc, ça me permet de rebondir sur le web sémantique, parce que je suppose que c'est la même chose pour cette partie-là, justement, vous avez demandé au prestataire un accompagnement, mais il n'y a rien de précis...

Ah, non, et le web sémantique, c'est ce qu'on a appelé un lot complémentaire, en fait. Voilà, comme je le disais, il va falloir qu'il nous prouve ce qu'il est capable de nous amener avec, et de la même manière on travaillera avec lui, je ne sais pas ce qu'on pourrait faire mais une fois que le premier lot sera sorti, on pourra envisager...

Donc, il n'y a aucune piste qui est envisagée pour l'instant ?

Pas vraiment non. On ne savait pas avec quoi partir exactement, et puis c'est toujours pareil on marche sur des budgets, donc entre la refonte d'un site, comment on le veut, plus le web sémantique, ça peut vite monter, on est tenus par le budget. Ce qui veut dire qu'une fois que le

premier lot soit sorti, on pourra demander de nouveaux budgets, coup par coup, pour aller vers le web sémantique. Si les budgets sont acceptés, alors, avec le titulaire du marché, alors on pourra partir sur cette voie-là.

Et justement, pour tout ce qui est marchés, comment ça se passe ?

Alors, pour les marchés... Donc les marchés sont diffusés sur une plateforme pour que toutes les sociétés puissent répondre. On est parti sur un appel de l'offre, donc à partir de l'appel d'offre... c'est-à-dire qu'au départ on n'a pas un budget limité, et donc ensuite on a fait un appel de MAPA, où là on a un plafond, qui est de 250 000 euros pour celui-là. Donc, on part sur un marché de quatre ans, justement, ouvert. On est parti sur cette idée justement en se disant « on ne sait pas trop où on va par rapport au web sémantique et tout ça, donc on a un budget de 250 000 euros TTC, et on sait très bien que si on part sur quatre ans avec du web sémantique, on va le dépasser. Donc, ils ont un mois, à compter d'aujourd'hui, pour répondre à l'appel d'offre. Nous, nous avons un mois pour découvrir l'offre. Ensuite, tous les prétendants au marché vont devoir passer en audition et ensuite, une fois qu'on a le grand gagnant, il est engagé avec nous sur quatre ans dans le projet.

Mais, au niveau des critères, comment on les détermine ?

Alors, les critères, il faut que je les établisse, il faut que je prépare une grille.

Et tu dis "je", du coup c'est toi qui es chargé de cette chose-là ?

Oui, alors je travaille avec le service de la commande publique, qui est un soutien par rapport à ça, parce que nous on ne fait que rédiger le CCTP, donc le cahier des charges, et eux se chargent de la partie administrative. Après, moi je vais leur proposer une grille de validation, par rapport à ce qui est attendu, je sais à peu près vers quoi on va aller, mais on verra. Et on va être deux ou trois à dépouiller les offres, à mettre des notes par rapport aux critères. On a une note qui a 35 %, déjà, donc celui qui a les points maximums remporte le marché, et après par rapport à comment est détaillée leur réponse à ces questions, c'est comme ça qu'on arrive à déterminer un classement. Là, vu que j'avais découpé en sept catégories le cahier des charges, par rapport aux attentes, on va faire cette grille, ils vont commencer à répondre. La première prestation qu'on demande c'est la mise en route du projet, justement de manière agile, qui va permettre de juger comment le prestataire peut nous accompagner, et délimiter...

Donc, comme on demande une méthode agile, on sait qu'il va y avoir plein de structures qui vont nous répondre et qui n'ont pas les épaules pour le faire. Donc il va falloir qu'ils soient capables de démontrer que, pendant quatre ans, ils seront capables de nous accompagner.

D'accord, et justement, cette volonté de passer par des méthodes agiles, ça vient d'où, qui a impulsé ça ?

Alors, l'agilité, elle a deux ou trois ans au Département.

Donc, c'est une volonté globale ?

Globale, oui, et qui dans le contexte, finalement, n'est pas trop mal, puisque justement on n'arrive pas à cibler vraiment... on n'est pas capables de faire un cahier des charges pour dire « il faudra ça, ça, ça et ça ». Donc finalement faire quelque chose d'un peu plus ouvert, ça peut avoir un coût bien supérieur, puisqu'on part du jour/homme. Donc si tu achètes 20 jours/homme, ce n'est pas comme si tu achetais un forfait, ce n'est pas la même chose. La méthode veut qu'à la fin on voit si on y est arrivé ou pas et à la fin à ça un coût quoi...

Annexe 7e : transcription de l'entretien avec Sylvain Thénault

Entretien avec Sylvain THENAULT, responsable de l'agence de Logilab à Toulouse et responsable du projet SAEM chez Logilab (prestataire ayant développé le référentiel du SAEM et le portail data.bnf.fr)

Entretien réalisé en visioconférence (via appear.in) et enregistré avec Audacity, le 13 juillet 2017 à 14h15.

Déjà, pour contextualiser l'entretien, est-ce que tu pourrais me parler un peu de toi, de ton parcours, de ce que tu fais à Logilab ?

Ouais, alors je ne sais pas quel niveau de détail tu veux...

Brièvement. *rires*

Dans les grandes lignes ?

Dans les grandes lignes, oui.

Alors, je suis ingénieur chez Logilab depuis mon diplôme en 2001/2002. J'ai participé à pas mal de projets, dans le cadre de Logilab, notamment autour du web sémantique. J'ai pas mal architecturé le logiciel CubicWeb, qui est sous-jacent au référentiel. Et sinon, je m'occupe de l'agence de Logilab à Toulouse. Et c'est aussi moi qui ai géré la partie, on va dire administrative, et une bonne part de la technique, du référentiel, avec l'équipe de Bordeaux.

D'accord.

Du coup, en termes d'interactions avec les archivistes, qu'est-ce que ça t'a amené à développer comme contact avec eux, dans le cadre du développement du référentiel ?

Alors, je ne sais pas si ça va répondre à ta question, mais je vais essayer. Donc, le projet de référentiel, c'est un projet qui se voulait en développement en mode agile. Donc, ils avaient quand même bien travaillé dessus, il y avait un cahier des charges dans lequel on sentait qu'il y avait pas mal de travail de fait, mais en même temps c'était un projet conséquent, ils ne savaient pas forcément trop par où prendre le problème. Nous non plus, d'ailleurs. Du coup, on a mis en place un certain nombre de pratiques d'agilité, notamment un type de gestion de projet type Scrum, avec itérations courtes. On se voyait, à l'époque... Nous, on allait à Bordeaux tous les mois, à peu près, on va dire pour simplifier qu'on faisait une itération par mois. Donc on faisait trois semaines de code, on leur livrait quelque chose, ils le testaient, ils mettaient des choses dans l'extranet et à la fin on se faisait une réunion en présentiel pendant laquelle on présentait tout et on faisait les retours et on décidait de ce qu'on faisait la fois d'après.

D'accord, donc dans le cahier des charges initial, il n'y avait pas, finalement, grand-chose, entre guillemets ? C'est une vision qui s'est développée au fur et à mesure ?

Ouais, enfin il y avait déjà... il y avait un peu leur vision, ça représentait un peu leur vision à l'époque, qui n'est pas forcément celle qu'on a faite. En tout cas, il y a toute une partie qui n'a jamais été vraiment faite, pas forcément parce que ça ne les intéressait pas mais parce que... En gros, dans le cahier des charges, il y avait deux projets différents : le référentiel et un autre. Les priorisations ont fait qu'on a mis du temps sur le référentiel et pas sur l'autre. Parmi les choses qu'on peut mentionner aussi, c'est qu'au début du projet j'ai fait une formation autour du web sémantique à toute l'équipe du projet. Et le dernier truc que je pourrais te dire aussi, c'est qu'il y a quand même eu une grosse évolution de l'équipe entre le départ et la suite, ça a augmenté en maturité et une fois que... Du coup c'était quand même relativement stable côté Bordeaux et du côté chez nous, l'équipe, ce qui a permis à toute l'équipe de Bordeaux de monter en compétences sur cette notion d'agilité, sur l'efficacité des réunions, sur les notions de priorisation et tout ça, et

qui a rendu la chose assez fluide, en tout cas beaucoup plus fluide que ce qu'on a dans nos rapports avec les clients.

Et quand tu as fait ces formations-là, est-ce que tu as ressenti que les différentes cultures professionnelles des différents acteurs du comité projet du SAEM compliquaient les échanges ou pas ? Ou est-ce que finalement ils ont développé une vision commune rapidement ?

Alors, est-ce que ça a compliqué les échanges... Clairement, il y a des gens qui avaient différents points de vue sur les aspects techniques et qui du coup ne partent pas tous du même niveau, il y a des choses qui sont plus ou moins faciles à appréhender pour certains que pour d'autres, et ça se voyait aussi dans les réunions de gestion de projet, de manière générale, pas que sur les choses techniques, mais également fonctionnelles. Ils n'avaient pas forcément les mêmes intérêts, on va dire. Typiquement, entre nous qui avons une certaine vision, Pascal qui avait aussi une vision un peu plus "métier" mais aussi assez technique et les archivistes qui s'intéressaient à des sujets assez précis, pour moi oui c'était le cas et c'était normal. C'était aussi l'intérêt de cette équipe projet que d'avoir ces différents profils.

Et je sais que chez Logilab, vous avez travaillé sur des projets similaires, en tout cas des projets dans le domaine de la gestion documentaire au sens large, par exemple avec la BnF. Est-ce que tu peux m'en parler un peu plus, surtout sur comment se déroulent ces projets ? Tu disais justement qu'en termes de gestion de projets, à Bordeaux, ça s'est mieux passé...

Oui, ça s'est mieux passé qu'avec la plupart des clients. Dans le cas de la BnF, je vais avoir du mal, parce que ça n'est pas moi qui ai géré ce projet-là. La vision que j'en ai, c'est qu'à la BnF, il y a quand même une démarche assez agile, en tout cas dans l'esprit. Même si l'une des grosses différences, quand même, avec la BnF ou si je prends l'exemple du portail France Archives, une réalisation un peu plus récente que je n'ai pas pu gérer donc je n'ai pas beaucoup plus à dire, en gros ces deux projets là c'est du forfait dans lequel on essaie de gérer le développement en agile. Je ne sais pas si c'est clair pour toi mais du forfait, ça veut dire que nous, en tant que Logilab, on s'engage sur une enveloppe de fonctionnalités. C'est-à-dire qu'eux ne vont nous payer que quand on aura fait tout ça. Dans le cadre du référentiel, on avait une vraie démarche agile, où on ne s'engageait pas de notre côté à faire des choses particulières. Alors, évidemment, on s'engage quand même sur un résultat, en gros notre engagement c'est que le client soit content, mais il n'y avait pas de... on travaillait au temps payé, quoi. Et du coup, c'est ça aussi qui a fait que le projet a été amené vers quelque chose qui n'est pas forcément ce qui était spécifié au départ.

Et justement, tu dis que ça pas mal évolué depuis le début, dans quelle mesure est-ce que ça a évolué, sur quel plan est-ce que tu sens qu'il y a eu des modifications, et qu'est-ce qui peut expliquer ces changements ?

soupir Euh, ce qui peut expliquer... Je n'en sais rien, c'est difficile à dire. Encore une fois, je pense que la vision de départ était trop floue et couvrait des périmètres trop larges pour pouvoir dire... En fait, il y a une part de sous-ensemblisme qui était indispensable pour pouvoir être... C'est-à-dire qu'il n'y avait pas, enfin je pense que personne n'avait, enfin nous on n'avait pas la volonté, et eux peut-être pas le budget non plus de faire tout ce qu'ils voulaient, donc il fallait faire des choix. Ça, c'est une part des choses, et il y a aussi pas mal de choses qui sont liées aux opportunités ou aux décisions qu'on peut prendre au jour le jour. Typiquement, à un moment donné, ils ont eu – alors, je ne sais pas comment c'est venu – mais il y a eu l'histoire d'interaction avec Arkhéia, un logiciel de gestion d'archives, qui n'était pas du tout prévue au départ et à un moment donné on a fait des choses pour ça. Il y a eu des choses comme ça, qui n'étaient pas dans la vision de départ. Donc si tu veux, au départ, il y a la vision d'un truc qui est gros, au fur et à mesure que tu pars de quelque chose de petit, mais tu vas dans l'objectif de montrer quelque chose, l'idée c'est de se focaliser sur ce qui est nécessaire. Et à la fin tu arrives nécessairement sur un truc qui est

sensiblement différent de ce à quoi t'avais pensé au départ, avec pour bonne partie des choses en moins, parce que finalement tu te rends compte, ou pas... enfin, c'était des choses auxquelles t'avais pensé, tu t'étais dit que ce serait bien mais rien ne te prouvait que tu en avais vraiment besoin, donc pas mal de choses en moins, et pour quelques trucs, des choses en plus, parce que tu n'avais pas pensé que tu pouvais avoir besoin de ce module d'échange avec Arkhéïa, par exemple.

Donc, ça veut dire que vous avez dû travailler conjointement avec la société qui développe Arkhéïa, Anaphore ?

Oui oui, enfin conjointement plus ou moins, mais il a fallu qu'on se parle un peu.

Et justement, ce dialogue-là, étant donné qu'eux ont une vision très documentaire de l'informatique, peut-être, comment est-ce que ça s'est mis en place et déroulé, globalement ?

Euh...

Ils ont compris les attentes, et vous, vous avez vu les contraintes liées à leur outil à eux, aussi ou... ?

En termes d'attentes, ce n'est pas vraiment passé par nous, c'était directement le COPROJ, c'est l'équipe de Bordeaux qui a fait ça. Après, du coup, ils avaient plus ou moins compris l'aspect technique. Après, c'était globalement des techniques qu'ils ne connaissaient pas trop... Nous, on n'a pas ressenti non plus une grande volonté de les connaître et du coup on leur a fait un peu de support et on leur a livré un truc qui leur minimisait complètement le degré d'interaction avec notre outil, si tu veux, c'était un peu compliqué. Et du coup, on leur a filé... en gros leur environnement c'était du Windows, on leur a filé un exécutable Windows qu'ils n'avaient qu'à appeler et qui faisait ce qu'ils avaient besoin de faire, quoi. Après, pour la partie web sémantique, pour eux c'était relativement intégré, par exemple les vocabulaires SKOS, des trucs comme ça...

C'est des choses qu'ils avaient déjà vues avant ?

Oui, dans le projet Arkhéïa il y avait déjà l'import SKOS.

Et justement, sur les technologies du web sémantique, globalement, comment est-ce que toi, tu mets en avant les avantages de ces technologies-là pour les produits que vous éditez, que vous concevez ?

Alors, qu'est-ce que tu entends par les technologies du web sémantique, exactement, est-ce que c'est les normes du W3C, ou c'est autre chose ?

Oui c'est ça, par exemple, les normes du W3C, mettre en place des vocabulaires avec le SKOS, des choses comme ça...

Bah, l'avantage se résume, pour moi, en un mot, on va dire principal, c'est l'interopérabilité. Donc être capable de faire des systèmes qui marchent ensemble plutôt que la démarche opposée, à savoir faire un gros système qui retient tout, en tout cas pour ce qui est de la base du web sémantique. Aujourd'hui il y a tout un tas de trucs... ça va assez loin sur d'autres sujets, mais la base qui est : « on va parler avec des vocabulaires communs sur la base d'un protocole qui est aujourd'hui HTTP, en nommant les choses avec des URL », déjà rien qu'en respectant ces principes-là, qui sont finalement assez simples, on va assez loin. Après, il y a tout un tas de couches supplémentaires mais qui pour moi en tout cas, ne sont pas matures, ou pas... Oui, c'est plus du détail d'implémentation, après.

D'accord, et tu penses à quoi en particulier dans ces couches supplémentaires ?

Oh, là, pas forcément grand-chose, je n'en sais rien, mais par exemple tous les mécanismes d'inférence que tu vas avoir au niveau des schémas OWL, des trucs comme ça, des règles... Quand tu regardes l'espèce de dessin, où ils présentent les différentes briques normalisées par le W3C, c'est une espèce de carré avec des trucs qui s'emboîtent : RDF, RDFs, OWL et tous ces machins-là...

Donc pour l'instant, d'après toi, il y a une partie qui est superflue, finalement ?

En tout cas, qui est aujourd'hui assez loin d'être mature. Par contre, avec les quelques briques qui sont assez en bas, tu fais pas mal de trucs. Le gros du boulot, c'est de se mettre d'accord sur les vocabulaires et tout ça, c'est pas tant une question de technique qu'une question de travailler ensemble.

Et justement, le potentiel de cette première couche là, on va dire, est-ce que toi, tu as ressenti qu'il était perçu pour le comité projet du projet SAEM, et en particulier pour les archivistes ? Et dans quelle mesure ?

Je pense qu'il était... Alors, déjà, cet aspect-là faisait partie du cahier des charges, notamment Pascal, qui a pas mal œuvré derrière, je pense, pour que ce soit le cas. Donc, je pense qu'il avait déjà un peu prêché cette cause-là avant et du coup nous on n'a pas trop eu à convaincre les gens que ça serait bien de le faire comme ça. On est arrivé, ça faisait partie des choses qui étaient acquises. Après, est-ce que c'est perçu dans la solution qu'on a aujourd'hui... Oui, je pense que ça commence à l'être, si on commence à regarder dans le détail les interactions entre les différentes applications. On essaie par exemple d'avoir des identifiants, par exemple, qui sont globaux, et qui dépassent le cadre d'une seule application, ça permet de commencer à faire des choses, il y a des trucs qui se parlent entre eux et qui savent de quoi ils parlent. Et ça je pense qu'aujourd'hui c'est assez visible. Après, ça pourrait l'être sans doute plus, et je pense que c'est un effet de bord technique qui n'est pas forcément... Enfin, je sais pas, il faudrait leur poser la question à eux, s'ils voient ça comme un truc indispensable ou pas... *rires*

Donc, à travers de ce que toi tu connais du cahier des charges et de la manière dont ça a évolué au fil du temps, est-ce que tu... Enfin, je ne sais pas si tu as la proximité, entre guillemets, nécessaire, pour répondre à cette question, mais est-ce que tu vois à quoi ça va servir, je veux dire au niveau concret, pratique, pour les archivistes, et éventuellement même pour les publics de ce genre de structures ?

Alors, j'ai une idée de ça, mais je n'en ai que l'idée, puisque pour moi, ça fait partie des... je ne sais pas si je dois dire échecs, mais en tout cas des problèmes potentiels, que nous n'avons aujourd'hui toujours pas réglé, c'est que normalement dans un projet agile... Alors on a bien réussi l'aspect... on va dire l'aspect humain entre les équipes de dev et les équipes projet mais par contre on a raté un truc, c'est que ça aurait dû être plus ou moins déployé, entre les mains des utilisateurs finaux, depuis longtemps. C'est un problème, aujourd'hui, puisque les pseudo-utilisateurs finaux ne sont plus du tout représentatifs puisque c'est les membres de l'équipe projet qui sont sur-formés sur le sujet et sur les technologies sous-jacentes. Donc aujourd'hui, j'ai une idée sur comment ça va servir, j'ai pas trop d'idée sur à quel point ça répond aux besoins. Parce que pour moi, on a raté... enfin, on aurait dû avoir des interactions qui n'ont pas eu lieu.

Quand tu dis « les interactions », tu veux dire avec les vrais utilisateurs, avec des...

Oui, avec les gens qui sont censés l'utiliser à la fin... Ça ne se limite pas à Delphine, Audrey ou je ne sais pas qui, ça devrait être à des gens qui, potentiellement, vont arriver dessus, ne vont pas trop savoir, vont vouloir l'utiliser, ils vont avoir des choses qu'ils faisaient au quotidien qui vont maintenant être faites par là, etc. Moi, aujourd'hui, je ne sais pas complètement dire ne serait-ce que qui est le public exactement derrière tout ça et qu'est-ce qu'attend ce public. Ça c'est une chose, et l'autre chose c'est que comparé à ce qui était spécifié dans le cahier des charges, enfin de mon point de vue, avec la réduction du périmètre en l'état, il y a tout un tas de publics qui étaient ciblés au départ qui ne sont plus des cibles aujourd'hui, notamment... Aujourd'hui, ça se destine aux archivistes, alors qu'au départ on pouvait voir des volets avec des utilisateurs plus grand public, par exemple.

Et donc, le potentiel, à travers l'interconnexion avec le SIA, avec Arkhéïa, et éventuellement avec la prochaine version du site des Archives départementales, ça non plus tu ne le vois pas

vraiment ?

Non. Enfin, non... Tu parles de quoi, la connexion avec...

Avec Arkhéïa et éventuellement avec la manière dont ça va ressortir à travers la prochaine version du site des Archives départementales, qui veulent justement se baser sur le référentiel pour leur indexation et pour les fonctionnalités de recherche de leur futur site.

Non, ça par exemple, je n'ai pas de mal à le croire, enfin à envisager comment ça pourrait marcher, effectivement ça paraît logique. Mais aujourd'hui, je n'envisage pas cet usage-là, enfin je n'en ai pas connaissance.

Tu ne l'envisages pas parce que tu n'es pas forcément "placé" au bon endroit, ou c'est parce que...

Les deux, nous aujourd'hui on a fait notre truc qui a sa place au sein de l'architecture avec Alfresco, as@lae. Arkhéïa, ça été fait un peu... à un moment, ça n'a pas demandé grand-chose... Mais les usages que ça leur permet de faire, je ne sais pas, je ne les connais pas.

Là aussi, je ne sais pas si c'est une question à laquelle tu vas pouvoir répondre, mais est-ce que tu perçois l'intérêt, justement, de la mise en relation de vocabulaires d'horizons divers à travers les technologies du web sémantique pour l'indexation des documents archives, ou est-ce que c'est trop éloigné de tes préoccupations ?

Oui, oui, pour moi ça fait partie des trucs qui n'était pas discutable dès le départ, l'intérêt du web sémantique sur ce genre de solution.

Alors, là je ne suis pas bien sûr parce que je n'ai pas vu de document qui allait dans ce sens-là, de ce que j'ai vu du projet je veux dire, mais est-ce que tu peux me parler du... Je ne sais pas dans quelle mesure vous, vous êtes intervenus à ce niveau-là... de comment s'est organisée la conception du modèle de données pour le référentiel ?

Alors, on est carrément intervenus à ce niveau-là, c'était une grosse partie de notre boulot finalement. Justement, ce modèle était de mon point de vue pas clair, donc on est partis sur... ça se faisait dans les fameuses réunions dont je t'ai parlé, donc en gros en partant à la base d'un périmètre XML, nous on implémentait un modèle de données qui permettait de représenter... enfin de servir ce modèle. Et au fur et à mesure des nouvelles fonctionnalités, ça a évolué, avec quelques tournants assez importants, comme par exemple, à un moment donné, on n'avait qu'une seule classe, je ne sais pas quelle terminologie tu préfères, pour représenter les notices d'autorité EAC et les utilisateurs instantanés du système, typiquement les archivistes qui se connectent au système, et au bout d'un moment on s'est rendu compte que ça n'allait pas. Ça, c'est surtout des problèmes qui ont été relevés par nous. Ils nous demandaient des évolutions plus ou moins grosses et du coup ça menait à des discussions en atelier avec eux pour voir la direction qui était raisonnable de prendre, aussi bien d'un point de vue technique que d'un point de vue fonctionnel. Ça faisait partie de notre boulot, ouais.

Et ce modèle-là, il est toujours en évolution, finalement ?

Ah, bah un modèle de données ça évolue toujours...

Oui, bien sûr, mais...

Plus ou moins lentement ou plus ou moins rapidement, selon l'étape du projet.

Et donc en ce moment, je suppose qu'il n'y a pas eu beaucoup d'ateliers de...

Non, depuis six mois, on vient moins souvent, oui. Parce qu'on fait moins de trucs, parce que le rythme, du coup... Comme là, ils sont en attente de mise en production, nous on fait un peu de support mais c'est vrai qu'on ne fait pas grand-chose.

Et, bon je passe un peu du coq à l'âne, mais est-ce que ton travail de formation, ou juste tes interactions en général avec le COPROJ, t'a permis de voir des rapprochements entre les

procédures et la conception globale de l'archivage électronique entre archivistes et informaticiens... ou pas ?

Oui, je pense que c'est un truc sur lequel il y avait pas mal de convergence. Globalement, je trouve qu'on est arrivé à quelque chose... Enfin moi, je suis plutôt satisfait du projet, je pense que c'est aussi parce qu'il y a eu une forme de convergence technique entre les membres du COPROJ et nous. À la fois, eux convergent vers des aspects plus techniques et nous on convergeait vers des aspects plus fonctionnels. Après, comme je le mentionnais tout à l'heure, le vrai saut c'est quand on va passer à des archivistes plus grand public, qui ont moins l'approche technique que peuvent avoir tous les gens qui sont dans le COPROJ aujourd'hui. C'est plus cette étape-là, sinon il y a eu une convergence dans le sens où on a appris à travailler ensemble, à comprendre chacun nos contraintes et du coup à partager une vision commune de ces problèmes, oui.

Et est-ce que toi, et ton équipe, vous avez suivi des formations sur les aspects liés à l'archivage électronique ou pas du tout ? Vous avez approfondi ça au fur et à mesure du projet ?

Oui, c'est ça, oui. Du coup, il y a forcément une étape où on répond au cahier des charges, où on se renseigne un petit peu. Après c'est clair qu'au début, quand on faisait les réunions, il y avait des trucs qui étaient assez connus chez Logilab, on va dire. Je l'ai vu par ce biais-là.

Dès le départ ? Parce que tu m'as dit que tu étais chez Logilab depuis 2001...

Oui, oui, parce que le web sémantique, ça existe depuis... Enfin, les bases sont posées depuis la fin des années 90 donc... Mais assez rapidement, c'est des trucs sur lesquels Nicolas, le PDG de Logilab, avait une espèce de vision, sur ce que ça pouvait apporter, et donc ça a infusé par ce biais-là.

Je me posais aussi la question de la mise en œuvre des marchés, avec les collectivités et tout ça. Je me demandais comment est-ce que vous, en tant que prestataire, vous appréhendez ça et qu'est-ce que vous mettez en avant, et quels sont les critères que vous essayez de respecter, pour mettre en valeur vos solutions ?

Pour obtenir les marchés, tu veux dire ?

Oui, c'est ça, exactement.

Bah... Nous, ce qui est clair, c'est que dans un projet comme ça, en l'occurrence, ça dépend, il n'y a pas de réponse universelle mais sur ce projet-là, où il y avait une grosse partie de web sémantique en avant, quelque part on partait déjà avec... Enfin, il y a certainement quelques concurrents qui avaient sûrement les mêmes avantages mais la bonne réputation qu'on a et le fait qu'on ait fait Data BnF, qui était l'un des portails, quand même très leaders, voire le leader, de tout ce qui est web sémantique en France, et que ce soit quelque chose qu'on affiche comme étant un de nos dadas et en gros comme l'un de nos secteurs... de nos compétences. On s'est principalement appuyé là-dessus, je pense que c'est principalement ça qui a joué, plus après des critères de prix, etc. Mais sur les critères autres que le prix ou justifier de l'équipe ou des machins comme ça, enfin il faut regarder les différentes parties de la note mais en termes de notre différence par rapport à d'autres, je pense que c'est ça qui joue, ouais.

Annexe 7f : transcription de l'entretien avec Agnès Vatican

Entretien avec Agnès Vatican, directrice des Archives départementales de la Gironde et membre des COTECH et COPIL du projet SAEM.

Entretien réalisé dans son bureau, dans les locaux des AD cours Balguerrie-Stuttenberg, le mardi 18 juillet à 16 h.

Pour contextualiser l'entretien, est-ce que vous pouvez me parler de votre parcours, des études que vous avez suivies, de votre parcours professionnel ?

Oui, je suis Agnès Vatican, conservateur général du patrimoine. J'ai suivi un parcours classique : classe préparatoire, école des chartes, école du patrimoine. J'ai eu un premier poste aux archives départementales de la Somme, où j'étais adjointe de la directrice en charge du déploiement du logiciel métier, à l'époque qui s'appelait ARDENT, donc ça a été mes premiers pas dans l'informatique métier, on va dire, partie documentaire. Ensuite, je suis partie à l'étranger et quand je suis revenue, j'ai pris la direction des Archives municipales de Bordeaux, où j'ai réalisé l'informatisation du service, qui ne l'était pas, où j'ai lancé la création du site Internet et où j'ai conduit un important projet de nouveau bâtiment, qui a été inauguré un peu après mon départ, puisqu'en 2013 j'ai pris la direction des Archives départementales de la Gironde, où les projets numériques sont nombreux, comme vous le savez. J'avais, aux archives municipales de Bordeaux, déjà participé au groupe projet SAEM, en tant que ville de Bordeaux. Donc c'est un projet que je connais depuis l'origine, mais côté ville.

Et justement, ça fait combien de temps que vous êtes intégrée au COTECH et COPIL, du coup ?

Du coup, depuis le début. J'ai fait même partie des réunions préliminaires, mais à l'époque en tant que directrice des Archives municipales.

Et votre rôle au sein de ces comités-là, qu'est-ce que c'est exactement, qu'est-ce que ça sous-entend ?

C'est des moments de restitution, évidemment, par le COPROJ, de l'avancement des travaux. À un moment donné j'ai même été dans le comité projet, quand j'étais aux Archives municipales de Bordeaux, puisque la cheffe de projet étant absente, je l'avais remplacée pendant quatre mois, donc j'ai même fait partie du comité projet, du groupe projet. Donc, au COTECH et au COPIL, on est plutôt sur la stratégie, sur l'orientation du projet, dans un contexte partenarial, donc c'est un peu particulier, puisque ce sont des COTECH à chaque fois à l'échelon du partenariat, qui ne sont pas propres à la collectivité. Mon rôle, c'est celui-là effectivement, avec mes collègues bien sûr, de valider la stratégie, la feuille de route, voire parfois même demander des inflexions dans le projet, voilà. Et puis faire l'articulation, avec les élus bien sûr, puisqu'on a aussi ce rôle, les directeurs et les DGA, d'articulation avec le champ politique et voir comment on répond aux attentes des politiques mais aussi comment on les sensibilise, sur un sujet comme l'archivage électronique évidemment, il y a une sensibilisation nécessaire.

De votre point de vue, puisque vous avez été intégrée au COPROJ, comment est-ce que vous avez perçu les échanges entre les différents acteurs du COPROJ, les informaticiens, les archivistes ?

Ah, je trouve que ça a vraiment été une réussite de ce projet, même si aujourd'hui effectivement je suis ça de beaucoup plus loin, parce que, en fait, s'est constitué un groupe humain qui s'est engagé dans une piste un peu nouvelle, avec pour beaucoup une montée en compétences, voir une acculturation sur le sujet, et une qualité du travail intellectuel mené, intellectuel et technique, mené dans ce cadre-là. Je reste aujourd'hui persuadée encore que ce partenariat autour de l'ingénierie est un point fort de ce projet, par rapport à d'autres.

Donc vous n'avez pas ressenti que les différentes cultures professionnels ont pu compliquer les échanges ?

Moi, j'ai plutôt ressenti qu'elles les enrichissaient, puisqu'au départ il y avait des profils informaticiens, profils archivistes, qui n'exerçaient pas forcément dans les mêmes sphères. Bon, c'était plutôt des archivistes qui étaient sur de la collecte d'archives contemporaines, donc qui avaient quand même un peu tous cette orientation. Sur les informaticiens, on était sur des profils chef de projet, donc les directions informatiques sont organisées comme ça depuis longtemps. Donc pour eux, je dirais que ce fonctionnement en chef de projet était plus habituel, pour les archivistes c'était peut-être plus nouveau. Et puis il y avait une dimension juridique aussi, puisqu'il y a eu, pendant un temps, la responsable des archives de la CUB, qui était aussi juriste.

D'accord, et en termes de vocabulaire, de compréhension des aspects techniques, ce genre de choses, est-ce qu'il y a eu une vision commune qui s'est rapidement développée ?

Hmm, oui, oui, alors il y avait peut-être des gens un peu devant, comme Pascal Romain par exemple, mais malgré tout il y avait un assez bon équilibre. La problématique ça a été quand il y a eu des changements de chefs de projet, notamment côté DSI, ça a été le cas à plusieurs reprises, et je pense qu'en plus, au fur et à mesure où le projet a avancé, c'est plus difficile, peut-être, c'est une hypothèse, pour les nouveaux chefs de projet, d'intégrer... parce que la montée en compétences qui est attendue, elle doit être rapide. Quand c'était au début, voilà, tout le monde tâtonnait un petit peu, alors un nouveau chef de projet avait le temps de se familiariser avec le sujet, plus le projet a avancé, plus le rattrapage, en quelque sorte, était dur.

Donc, c'est plus sur l'aspect gestion de projet qu'il y a eu des adaptations nécessaires, plus que sur les échanges techniques entre les différents acteurs ?

Il me semble, après je ne suis plus au groupe projet depuis des années, mais je pense que le groupe projet a bien traversé des changements de chefs de projet, des réformes administratives avec la montée en puissance de la métropole, le départ de la région. Voilà, il me semble que tout ça a été traversé de façon correcte. La question qui se pose pour moi aujourd'hui c'est plutôt comment on passe de ce mode projet, très intégré, très performant, à la mise en production et à l'acculturation des équipes, parce que derrière, il s'agit bien que les équipes s'approprient l'outil et l'utilisent.

Et d'après vous, les équipes sont prêtes à intégrer ces changements, à découvrir ce nouvel outil ? Parce que la mise en production du système a été un peu repoussée, donc comment est-ce que vous percevez cette transition ?

Moi je pense qu'il y a encore un travail de conduite du changement à mener.

Et vous avez une idée de la manière dont ça doit se mettre en place, cette conduite du changement ?

Justement, on en parlait, moi je souhaite qu'il y ait une réunion à la rentrée, puis on avait prévu une information, une formation, mais c'est toujours compliqué de sortir du mode projet pour basculer dans la gestion quotidienne, or c'est bien vers ça qu'il faut qu'on aille. C'est-à-dire que l'archivage électronique ne pourra pas reposer sur une personne aux AD de la Gironde. J'exagère un peu mais voilà, on ne peut pas... Demain il faut que tous les cadres en charge d'archivage contemporain maîtrisent l'outil et soient capables de s'en servir.

Sinon, pour parler d'archivage électronique de manière plus globale, est-ce que pour vous ça a introduit une modification des pratiques, avec l'effort de numérisation des archives papier, ce genre de choses ? Et dans quelle mesure est-ce que ça a bouleversé les pratiques des archivistes ?

La numérisation, la montée en puissance de la numérisation ?

La numérisation, la gestion des métadonnées, la rédaction d'instruments de recherche

électroniques, etc.

Ah, il y a eu évidemment davantage de formalisme, davantage de normalisation, ça c'est inhérent, mais dès qu'il y a eu informatisation, même sur les aspects de gestion, sans parler de la description documentaire, quand vous avez une gestion de salle de lecture informatisée, il y a forcément quelque chose qui se met en place, qui est plus rigoureux, qui est plus normé. La montée en puissance de la numérisation, elle a produit surtout, je pense, des effets dans la relation avec le public, et en amont peut-être des déséquilibres internes entre les différentes missions d'un service d'archives, qui ont pu être bouleversées, notamment les activités de classement qui ont pu être reléguées au rang de... Enfin, reléguées, c'est un peu péjoratif, mais qui ont pu être perçues comme avant tout destinées à nourrir une chaîne de numérisation, alors que le volume de fonds numérisés reste extrêmement marginal, malgré des chiffres qui peuvent impressionner.

Vous parliez de normalisation, est-ce que toutes ces normes, tous ces standards, comme le modèle OAIS pour l'archivage électronique par exemple, est-ce que ce sont des choses qui se transposent aisément à la pratique existante pour l'archivage papier ? Est-ce que ça coïncide avec les pratiques, est-ce que les archivistes s'y font facilement ?

C'est plus moins évident, je pense que celle qui a le plus facilement, enfin facilement ce n'est pas vraiment le mot, mais qui est peut-être le plus entré en vigueur, c'est ISAD(G) et donc l'EAD. Après, sur le reste, je pense que les métiers de la documentation étaient peut-être plus disponibles, plus formés à ces modèles-là. De toute façon, la vision tout documentaire des archives où des métadonnées retireraient complètement la vision du fonds, dans son organisation, son plan de classement, ça c'est un peu la vision qu'il y a eu dès les années 70, parce qu'un logiciel comme ARDENT, dont je vous parlais, qui est une gestion essentiellement documentaire. Donc c'était de l'indexation, toutes les descriptions étaient à plat, et avec l'indexation on allait tout retrouver, ça c'est une vision d'esprit, en matière d'archives en tout cas.

Et justement, les instruments de recherche, dans leur forme actuelle, en tant que documents, ils constituent encore un outil pertinent ? Est-ce qu'ils sont adaptés aux usages des différentes catégories de publics, d'usagers ?

En fait, aujourd'hui un instrument de recherche c'est plusieurs niveaux. C'est-à-dire qu'au départ c'est effectivement un ensemble de métadonnées qui sont saisies de façon normalisée pour décrire un fonds, partons de ça par exemple. Ensuite, là où à mon sens, on a encore des progrès à faire, c'est effectivement dans les interfaces de consultation. Globalement, la profession des archivistes a été, tel que je le perçois... parce que dans le logiciel ARDENT, bon c'était les années 95-97, il y avait un groupe utilisateur qui était très actif et on était un nombre assez conséquent de départements, au moins une vingtaine, c'était des gens qui avaient été souvent pionniers de l'informatisation. Les archivistes ont été assez intéressés par ça, alors bien sûr ça dépend des profils, il y a ceux qui sont passés complètement à côté, mais voilà. Ensuite, ce qu'on a vu, ce qui est arrivé après, et notamment les sites Internet, mais c'était aussi le cas dans ces premiers logiciels de gestion : il y a les gens qui ont le point de vue du « tout métier » et qui vont défendre avant tout le « tout métier », et puis il y a des gens qui vont être plutôt sur la problématique de la réception par l'utilisateur, qui peut être l'archiviste ou qui peut être le public. Aujourd'hui, si on regarde un peu les mises en ligne telles qu'elles se font, on voit bien qu'il y a ces deux... Nous clairement, GAEL c'est du « tout métier », c'est-à-dire qu'on est dans quelque chose qui se veut très conforme à l'EAD. Bon, quand on va chercher dans les coins, il y a beaucoup de combinettes pour que les données arrivent à s'afficher, parce que quand on rétro-convertit des instruments de recherche du 19^e siècle, il est clair qu'on ne peut pas avoir un EAD propre, complet. Mais on a une interface qui est très orientée métier, et qui suscite de la part des usagers, on le sait parce qu'il y a eu une enquête sur l'ensemble des internautes, on n'avait un indice de satisfaction qui était en dessous de la moyenne des Archives départementales. Mais on est aussi sur un site qui est plus ancien, dans sa conception.

Et aujourd'hui, il y a un certain nombre de sites Internet, qui eux sont sur des interfaces de consultation des instruments de recherche et des fonds numérisés, qui sont beaucoup plus orientés publics, avec des accès très rapides à certains types de sources, notamment les fonds sériels. Donc clairement, on est entre ces deux cultures, ces deux visions. Et la question, c'est effectivement où on se situe, quel équilibre on trouve entre les attentes des uns et des autres. Mais un instrument de recherche en EAD, il peut être publié de plein de façons.

Vous parliez des sites d'archives, et je sais qu'il y a un projet de refonte du site, j'ai vu passé le cahier des charges, comment est-ce que vous percevez la place, le rôle, d'un site d'archives aujourd'hui ?

Aujourd'hui, clairement, on est arrivé à une maturité des sites d'archives, donc beaucoup de services comme le nôtre en sont à leur première, voire leur deuxième refonte. C'est un marché qui reste un marché de niche mais sur lequel il y a eu des évolutions importantes faites par les acteurs. On est sur cette double tension, avec des prestataires qui sont très orientés métier, d'autres qui sont très orientés publics, et à la fin ça fait des sites assez différents. Donc aujourd'hui, on voit bien, les internautes comparent, on n'est plus isolé, chacun dans son coin, l'information peut être trouvée ailleurs, avec la réutilisation, on sait qu'il y a un certain nombre d'acteurs qui vont se positionner, et que l'état civil et demain certainement d'autres fonds, seront consultables ailleurs que sur notre site Internet et puis on ne peut pas être non plus dans un vase clos. La question de la dissémination des données, elle est aujourd'hui stratégique, en tout cas elle fait partie de la stratégie numérique qu'on applique à notre projet scientifique et culturel. Je ne sais pas si vous l'avez lu ce document...

Je ne crois pas, non.

Ça serait bien que vous le lisiez, c'est le projet d'établissement, il a été voté au mois d'avril. Le PSCE, ça s'appelle, il y a toute une partie sur la stratégie numérique.

Et en parlant de stratégie numérique, quel a été l'impact de la feuille de route du ministère de la Culture concernant la transition web 3.0 et les métadonnées culturelles ?

Alors ici, il y a un intérêt manifeste pour ces questions-là. Après, nous on est même pas au web 2.0. C'est-à-dire qu'on a à la fois potentiellement l'aptitude à être pilote sur le web sémantique, avec des fortes expertises dans nos équipes. Ce qui est restitué au public pour l'instant, voilà... il n'y a même pas les outils du web 2.0, les pratiques collaboratives, tout ça n'est pas implémenté sur notre site, et c'est encore quelque chose qui nécessite un accompagnement au changement auprès des équipes ici. Par exemple, les pratiques collaboratives d'indexation, aujourd'hui c'est en standard dans toutes les offres de sites Internet, après vous en faites ce que vous voulez, vous en mettez ou pas, sur l'état civil ou autre chose, c'est une question de politique d'établissement. Ici, on a des niveaux d'expertise très pointus mais en réalité, si vous regardez notre site Internet, ce n'est pas quelque chose qui est restitué au public, en tout cas pour l'instant, non.

Et qu'est-ce qui explique cette restitution difficile au public ? Est-ce que c'est parce qu'il y a encore un retard parmi les différentes équipes, parce qu'il y a des spécialistes qui se sont appropriés ces questions mais ça n'a pas encore infusé dans les services ?

Il y a une problématique, oui, d'accompagnement au changement, clairement. Et puis, il y a aussi un projet de bâtiment qui a fait que sur le site Internet, il y a eu un investissement indéniable, mais c'est difficile, même pour un gros service comme le nôtre, de construire un bâtiment et en même temps... Voilà, donc on a un gros moteur de recherche, mais autour pas grand-chose, sur le site Internet, par rapport à ce qu'on voit ailleurs. Et l'interfaçage avec les métadonnées qui sont produites en interne pour le public, à mon sens ici c'est quelque chose qui n'a pas fait l'objet d'une réflexion assez poussée, donc c'est un peu l'objet, l'enjeu même, de la refonte : comment est-ce qu'à partir de ce travail énorme qui a été fait, de rétro-conversion, de production d'instruments de recherche électroniques normalisés, comment est-ce qu'on donne à voir cette richesse. Parce qu'aujourd'hui, à mon sens, le site Internet ne rend pas justice au travail qui a été fait. Il a une

interface de consultation qui est, on va dire en deçà du standard qu'on peut trouver sur d'autres sites, bon ça c'est vraiment un problème de calendrier, mais les questions sur les usages n'ont pas été assez... Et aujourd'hui on a la chance d'avoir eu cette enquête de 2013/2014 qui va nous permettre aussi... et puis la feuille de route du ministère, voilà, mais moi ce qui m'interroge beaucoup c'est les usages, je pense que c'est très important en fait. Parce qu'on peut avoir la description la plus conforme, à la pointe de tout ce qui est attendu en termes de métadonnées, mais si l'interface de restitution au public n'est pas opérante, ce travail-là n'est pas perçu par le public. GAEL, c'est facile, vous prenez un docteur en Histoire, il n'a aucun problème sur GAEL, vous prenez un usager qui est moins expert, il appelle au secours tôt ou tard.

Et d'après vous, est-ce qu'il faut différencier les différentes catégories d'utilisateurs directement dans les interfaces de consultation, est-ce qu'il faut prendre en compte ces différents niveaux ?

Ce n'est pas... De toute façon, je crois que faire simple c'est toujours compliqué. C'est plus compliqué que de faire compliqué. C'est comme expliquer simplement un concept très élaboré, c'est plus compliqué que de l'expliquer avec un niveau de complexité élevé. C'est ça aussi, la médiation, c'est un terme un peu galvaudé qu'on met à toutes les sauces, mais c'est effectivement arriver à traduire simplement une réalité complexe. Donc, ensuite par rapport à cette réalité complexe, parce que forcément les instruments de recherche dans un service d'archives, surtout avec des volumes de fonds comme les nôtres, c'est forcément quelque chose de complexe, qu'il puisse y avoir différents niveaux, voilà... Après, moi je ne suis pas favorable au fait de mettre les gens dans des cases et de dire d'entrée de jeu... C'est un peu le reflet, aussi, des salles de lecture. Au début de ma carrière, je me rappelle, il y avait des revendications quand il y a eu les généalogistes qui sont arrivés dans les salles de lecture, les usagers plus professionnels, universitaires ou autres, avaient la revendication d'avoir des salles de lecture spécifiques, avec des conditions d'accès spécifiques, etc. En fait, vous transposez ça sur le net et c'est effectivement... En réalité, on voit bien que l'utilisateur expert, quand l'interface est simple, il va aller vers cette ressource-là, lui aussi. Voilà, les universitaires aujourd'hui, vont beaucoup vers Gallica, qui est une interface assez simple de consultation, et elle n'est pas très complexe non plus pour le grand public. Même si Gallica, aujourd'hui, commence un peu à souffrir de son obésité, donc ça commence à devenir plus compliqué, donc ils font des sous-Gallica, des portails thématiques, etc., mais on voit bien que ces enjeux d'édition de contenus, à la BNF, il y a une grande attention qui est portée à cela. Et aujourd'hui, je crois que le principe du web, c'est ça, c'est vraiment offrir la simplicité à tous, y compris à ceux qui sont experts et qui pourraient... c'est un enjeu de démocratisation, certainement, qui est posé. Après, les réponses ne sont pas forcément uniformes et le modèle qu'on a eu souvent, d'une recherche simple et une recherche experte peut rester opérant, il ne faut pas tout jeter.

Et dans l'optique de faciliter l'accès pour les utilisateurs, comment est-ce que vous ici, aux Archives départementales de la Gironde, vous percevez, ou pas, le potentiel des technologies du web sémantique, de l'exposition des données issues du traitement des archives dans des grands portails, ce genre de choses ?

Eh bien je ne sais pas.

Oui ?

Oui, je ne sais pas. À un moment, la recherche à facette semblait être l'alpha et l'omega, aujourd'hui le web sémantique... Je ne sais pas.

D'accord, c'est encore trop flou ?

Oui, pour moi c'est encore trop flou... Je n'ai pas encore d'opinion trop tranchée sur ça, en tout cas je pense qu'il faut faire attention à ne pas penser que la technologie qui vient sera toujours la meilleure. Il y a un exemple, en ce moment dans la profession, qui nous amuse, c'est le portail FranceArchives, qui a été ouvert récemment, et comme aujourd'hui la politique du ministère de la Culture, c'est de ne pas multiplier les sites Internet, le SIAF a été contraint d'inclure dans ce grand

portail toute la documentation à destination des professionnels qu'on trouvait jusque-là sur le site des archives de France. Résultat : les professionnels des archives, qui avaient l'habitude, ne retrouvent pas les circulaires, les instructions de tri, etc. Et le SIAF vient d'envoyer un message assez long, une sorte d'aide pour nous expliquer comment retrouver les circulaires, qui jusque-là étaient accessibles avec un classement thématique qui était parfaitement opérant pour des professionnels. Là, on a effectivement un portail avec un grand champ d'interrogation, aujourd'hui c'est vraiment le modèle, y compris des sites institutionnels, puisque Gironde.fr vient d'être refondu sur ce modèle-là. Et donc vous tapez « tribunal de police », et là vous avez plein de trucs qui ressortent, et là ma circulaire, la dernière, actualisée, où est-elle ? Donc là, on a vu, quand même, les limites d'un portail où l'entrée ne se fait que par métadonnées. Alors, en bas il y a le plan du site, on peut retrouver un peu d'arborescence, s'y retrouver, mais ce n'est pas opérant. D'ailleurs, si vous allez sur le site de la BNF, vous comparez, vous avez Gallica, l'accès au catalogue, avec des modalités de navigation et de recherche de différents niveaux, et après vous avez des rubriques à disposition des professionnels, là on n'est pas sur du moteur de recherche, on est sur une navigation hiérarchisée et pour les professionnels que nous sommes, on sait où trouver l'information qui nous intéresse

Et donc, ça tombe bien, vous avez mentionné le portail FranceArchives, je voulais vous poser la question, mais est-ce que vous avez entendu parler d'autres projets, ressources, dans le domaine de la culture, qui reposent sur le web des données, sur cette mise en ligne de métadonnées et cette interconnexion ?

Euh, là comme ça... Non, comme ça, je ne vois pas, il y a quoi ?

Eh bien je pense par exemple, vous parliez de la BnF, Data BnF...

Ah oui !

Et je ne sais pas si vous êtes familière de son fonctionnement ou pas...

Non...

Non ? Ok.

Et pour revenir sur la description archivistique à proprement parler, est-ce que, justement, vous pensez qu'il y a une interopérabilité entre la description classique et cette tendance à aller à un niveau de granularité assez fin, à privilégier la donnée, ce genre de choses ?

Je pense que l'un ne doit pas exclure l'autre. Enfin pour moi, FranceArchives, c'est quand même assez révélateur de ça. Il y a la forme livre, qui peut être PDF, etc. La forme livre, elle reste opérante à bien des égards. Je me souviens d'une rencontre professionnelle, je crois que j'étais encore aux Archives municipales de Bordeaux, où un conservateur de la BNF m'avait repris par rapport à une présentation en me disant « mais enfin, toutes ces publications des sociétés savantes locales, elles sont en ligne sur Galica », je lui avais dit « oui, oui, ça je le sais bien mais dans la salle de lecture des Archives municipales, ces revues d'histoire locale, qui sont en fait des usuels, sont sous format papier dans la salle de lecture », parce qu'effectivement, quand on a en tête la référence de l'article, c'est tellement simple d'attraper le volume et de le feuilleter. Et puis, la consultation d'instruments de recherche, on va dire classiques, pour parler comme ça, ça permet aussi, alors c'est un sujet qui m'intéresse beaucoup, qui est la sérendipité, c'est-à-dire le fait de tomber sur quelque chose qu'on ne cherchait pas. Et ça, le web de données, je ne dis pas qu'il l'exclut, mais en tout cas il le rend moins probable.

Dans quelle mesure ? Enfin, comment est-ce que vous expliquez ça, qu'il y ait moins de sérendipité ?

Bien, parce que la métadonnée, elle enferme la recherche. Elle enferme autour d'un terme, autour d'une notion, beaucoup plus que le feuilletage, qu'il soit en ligne ou sur papier, qui permet ça. Voilà, l'historien italien Carlo Ginzburg raconte comment il cherchait un sujet de recherche, c'est un des papes de la microhistoire, donc il va aux archives à Venise, les archives de l'inquisition, il

passé des semaines à consulter des liasses de procès et diverses choses et à un moment donné il est tombé sur des procès d'une catégorie de personnes très particulière, ce n'est pas les loups garous mais bon voilà, des gens qui se transformaient en d'autres personnes et qui étaient pourchassés par l'inquisition, et il en fait un sujet de recherche absolument fabuleux. Il se trouve que j'ai travaillé sur des sources d'inquisition en Espagne et effectivement, le dépouillement des instruments de recherche était réduit à un petit registre photocopié de quelques pages, qui était censé donner accès à des fonds d'inquisition énormes, donc il n'y avait aucune description, aucune métadonnée quasiment, c'était le tribunal, l'année et voilà. Il n'y avait même pas le type de code, ni rien, il fallait dépouiller ce type de fonds comme d'autres l'avaient fait, ce que j'ai fait, et c'est vrai que le dépouillement ne peut pas être remplacé par une description par métadonnée. On passe à côté de certaines choses, même s'il y a toutes les possibilités de rebonds, etc, qui permettent d'aller d'une notion à l'autre, ça je le sais, je le pratique, j'aime d'ailleurs beaucoup le catalogue de la BNF pour ça, il permet vraiment des parcours de recherche. Mais je pense qu'il y a quand même toute une partie d'imprévu qui est moins certaine que quand on prend un inventaire et qu'on le lit de A à Z.

Du coup cette place du document, qui est finalement un peu sacrée dans les archives, est-ce que, de ce que vous avez pu observer, est-ce que c'est quelque chose qui constituerait, en termes de culture professionnelle je veux dire, une sorte d'obstacle à cette transition, peut-être perçue comme forcée ?

Eh bien ça ne peut pas être un obstacle dans la mesure où c'est la justification même de notre métier. Donc il faut arriver à trouver des articulations mais penser que tout deviendra donnée, à mon avis c'est une utopie un peu dangereuse.

J'aurais voulu avoir votre avis, je change un peu de sujet, sur la place du module référentiel dans le projet SAEM. Comment est-ce que, vous, vous envisagez sa place, ses enjeux ?

Alors, assurément c'est un point fort de ce projet, c'est un outil en plus qui pourra être réutilisable par d'autres, donc c'est très intéressant. J'ai eu deux démonstrations, je trouve l'objet encore un peu complexe, il parle surtout, je trouve, à ceux qui l'ont conçu, mais bon je n'ai pas eu l'occasion de m'y pencher plus attentivement. Il faudra voir comment ça va pas pouvoir être pris en main par les professionnels, ça pour moi c'est important. C'est une brique qui manquait, très certainement, donc c'est bien de l'avoir fait et la mise en production devrait nous permettre de valoriser ça, en tout cas pour moi c'est l'idée, c'est de valoriser ça dans la mise en production. En tout cas, ça signe un peu une caractéristique du SAEM. Après, il ne faut pas oublier que les attentes, elles sont quand même de pouvoir verser des flux, et il ne faut pas qu'on prenne de retard par rapport à ça, par rapport à d'autres projets qui avancent bien de leur côté.

Je ne sais pas dans quelle mesure vous avez suivi ça, mais vous avez entendu parler du modèle RiC, je suppose, Records in Context, pour la description, qui modifie un peu la manière dont les archives pourront être décrites.

Oui.

Quelle influence est-ce que ça pourrait avoir selon vous dans la pratique et est-ce que ça serait une mutation qui serait bien accueillie par les professionnels ?

Je pense que le fait d'être passé à l'EAD peut préparer aux étapes suivantes. À mon avis, le saut le plus important, c'était, voilà... Après, c'est comme tout, l'EAD, c'était bien passé dans nos services comme dans d'autres, parce que le logiciel qu'on avait l'embarquait à l'insu de l'utilisateur, c'est-à-dire qu'en fait vous pouvez n'avoir jamais fait d'EAD et malgré tout produire des instruments de recherche en EAD, bon ce n'est pas d'un niveau très complexe mais ça c'est, il faut ensuite que ces notions-là... Enfin, j'avais fait une formation il y a quelques années sur la DTD EAD, sur le XML, bon je veux dire, lire le XML en parfait bilingue, ça n'a quand même un intérêt très limité. Bon, il y a des gens qui vont développer une expertise très pointue, qui vont pouvoir lire couramment le

XML, corriger une balise à la volée, d'un simple coup d'œil etc, mais après il ne faut pas penser qu'on va embarquer toute une profession avec ce genre d'outils, clairement pas. Moi aujourd'hui, c'est ce que je reprocherais aux GED, c'est que c'est des interfaces qui sont extrêmement arides, donc on n'attrape pas les mouches avec du vinaigre. Les gens ont lâché leurs fiches cartonnées ou leurs répertoires Word parce qu'à un moment donné ils ont eu des outils simples où ils retrouvaient dans les intitulés les champs, et je pense que c'est la même chose en bibliothèque, ils ont lâché les fiches à tringles quand il y a eu les logiciels, voilà. Par contre, coder le MARC, en faisant "\$\$", etc., bon ça, ça n'a été réservé qu'à une toute petite partie et rapidement on est arrivé vers les dérivations de notices, l'exemplarisation, etc. Bon, on ne peut pas faire ça parce qu'on est sur des fonds, par définition, originaux et en général uniques, on peut le faire un peu sur les notices de producteurs même s'il faut à chaque fois les adapter au contexte local, mais je pense qu'il faut arrêter de penser que les professionnels peuvent trouver, en plus dans un contexte où les ressources en moyens sont comptées, le temps de se former, tous les 10 ans, à un nouveau modèle de données. Il faut qu'il y ait de la médiation, il faut qu'il y ait de l'accompagnement et il faut surtout qu'il y ait des outils qui rendent l'exercice simple. Et puis je dirais aussi qu'il faut aussi veiller, comment dire... Il y a eu une époque, aujourd'hui c'est beaucoup moins le cas, où on faisait évoluer le cadre de classement réglementaire et donc tout le monde se mettait à recoter, donc on ne finissait jamais non plus les chantiers parce que... des fonds qui étaient cotés, qui avaient des instruments de recherche, d'un coup la circulaire changeait, « pom pom pom », on décotait, on recotait, les tables de concordance, puis entre temps on n'avait pas le temps de finir, voilà. Moyennant ça, on n'a encore des fonds qui ne sont pas classés, qui n'ont pas d'instrument de recherche, donc se dire : tout ce qui était en EAD, on va le repasser dans un nouveau modèle de description etc, on va tout reprendre, alors qu'on a des fonds par ailleurs qui n'ont pas d'instruments de recherche... Moi, je pense qu'il vaut mieux temporiser, c'est un peu comme les gens qui veulent toujours avoir la dernière voiture ou le dernier... Voilà, c'est quand même un peu une course perpétuelle, on est des institutions qui nous inscrivons dans une certaine pérennité, on a deux siècles derrière nous de fonctionnement, il y a encore des instruments de recherche du 19e siècle qui sont parfaitement opérants et qui permettent un accès aux fonds. Donc, je pense que vouloir, à marche forcée, basculer de modèle en modèle de description... Je pense qu'il faut faire attention, et considérer une stratégie plus globale, qui est quand même l'accès aux fonds, ce qui est important c'est de donner accès à nos fonds, parce que le lecteur derrière, l'EAD, Records in Context, etc., il ne connaît pas, il n'a pas à connaître et il ne veut pas connaître. Par contre ce qui l'intéresse, c'est de savoir ce qu'on a. Donc avant de se dire qu'on va tout rétro-convertir dans le nouveau modèle de description, il faut quand même s'interroger sur la stratégie qu'on doit développer. Après, on verra bien, je pense que les outils s'adapteront, intégreront ce nouveau modèle, proposeront des solutions pour un travail ergonomique. Et là, voilà, pourquoi pas. Mais je me souviens, sur le logiciel ARDENT, la description documentaire était extrêmement complexe, la saisie, entrer deux termes d'indexation, des thésaurus qui n'étaient pas chargés, enfin bon... Je pense qu'aujourd'hui, où l'administration est de plus en plus sur une recherche de performance, et qui dit performance dit outils adaptés et aussi dans un certain nombre de cas continuité dans le fonctionnement.

Vous m'avez dit avoir suivi des formations sur l'EAD, le langage XML, et globalement, vous avez suivi quel genre de formation sur ces aspects-là, sur l'archivage l'électronique, sur la mise en ligne des documents d'archives, ce genre de choses ?

En formation initiale, il y avait assez peu de choses, on avait quand même à l'école des Chartes une formation aux langages informatiques, comment ça s'appelait... le DOS, oui le DOS, ça m'a servi une fois je me souviens, je me suis trouvée avec un gros plantage informatique sur mon portable, j'étais au fin fond de l'Espagne et j'ai réussi à aller récupérer des fichiers avec des commandes du DOS. Après, oui, il y a eu des choses sur la numérisation, mais c'était encore très émergent, moi c'était plutôt sur le terrain que, dès mon premier poste, on avait commencé à numériser. Alors

c'était très embryonnaire, il n'y avait pas encore de site Internet, ça n'existait pas. Moi j'ai fait quelques formations, le fait d'avoir conduit deux informatisations, en quelque sorte, c'est quand même quelque chose qui permet d'apprendre pas mal. Après, c'est surtout sur l'EAD, j'avais fait une formation sur l'encodage DTD EAD et TEI, il y avait les deux d'ailleurs, les principes du XML, de l'encodage, les outils etc, on avait quelques petits travaux pratiques. Et j'avais fait aussi une formation sur la rétro-conversion des instruments de recherche. Sur la numérisation, je ne me souviens plus très bien, c'est une question sur laquelle je ne me suis pas sentie... je me suis sentie plutôt à l'aise.

Alors, j'avais noté que vous aviez mentionné le rapport Nougaret pendant le COTECH.

Oui, vous l'avez lu ?

En partie, oui, du coup je me demandais dans quelle mesure les éléments qu'on y retrouve coïncident avec par exemple ce qu'il y avait dans la feuille de route du ministère de la Culture sur la transition web 3.0, avec justement ce dont on parlait tout à l'heure ?

Eh bien, il se situe à un niveau plus élevé, en fait. C'est un peu sur des enjeux, des grands enjeux. Et puis c'est un rapport qui a été fait très rapidement, mais ça c'était, me semble-t-il, une volonté du ministère, d'avoir très rapidement un état des questions, et ce qui est intéressant surtout, c'est qu'ils étaient partis, effectivement, de la question de l'accès, et puis en réalité il est remonté, et c'est devenu la partie la plus prépondérante, sur la collecte et puis la notion de fonds d'archives. Donc sur le numérique, il pose un certain nombre de questions, effectivement des problématiques de ruptures entre le producteur et le territoire, la notion d'original, la question de la sélection qui se pose plus que jamais, effectivement, face au volume de données nativement numériques.

Mais par exemple, on n'y retrouve pas du tout les notions, dans la partie dédiée à l'accès, de web des données, de web sémantique.

Non, mais je crois que ce n'était pas ça le... Ça existe dans d'autres documents, là c'était plutôt des enjeux un peu plus stratégiques. On est plutôt dans les usages, dans les offres de services. Et puis bon, elle a pointé quand même, d'ailleurs ça va apparemment faire l'objet d'un travail, les problématiques de législation et de réglementation de l'accès, c'est un véritable maquis, avec la codification. Donc ça c'est certainement quelque chose qui va se faire, après il y a des propositions de portées et de natures différentes, dans ce rapport. Après, ce n'est pas un rapport qui se veut... C'est « au regard des évolutions, quels sont les enjeux aujourd'hui pour les services d'archives ? », ça s'adresse bien aux services d'archives, d'ailleurs ce sont essentiellement des archivistes qu'elle a rencontrés pour élaborer ce rapport. Après il y a des choses, comme promouvoir le collaboratif avec les internautes, assurer la présence institutionnelle sur les réseaux sociaux, aussi, même si les Archives départementales de la Gironde n'y sont pas encore, mais on voit que c'est quand même... diversifier les offres de service aux utilisateurs distants, une médiation en ligne. Voilà, tout ça c'est des choses qui sont déjà, je dirais, dans la feuille de route, de la plupart des services d'archives. Par contre, les enjeux sur la législation et la réglementation de l'accès, c'est peut-être moins connu du grand public mais effectivement, il y a un maquis, une opacité, dont il faut absolument sortir, parce qu'en plus on voit que chaque nouvelle loi qui est promulguée vient rajouter la complexité, des régimes spéciaux...

Et est-ce que les questions d'open data, d'ouverture des données rajoutent une couche de complexité pour les services d'archives ?

Ah oui, oui, clairement. Voilà, c'est une couche de plus, c'est un régime de plus, effectivement, d'accès aux données, bien sûr.

Le rapport rappelle aussi, par rapport à la relation aux services producteurs, l'importance de se placer en amont pour l'archiviste, dès la création du document, bon ce n'est pas quelque chose de nouveau, mais est-ce que vous diriez qu'il y a de plus en plus de convergence entre les approches de l'archivistique traditionnelle, on va dire, et de ce qu'on appelle le records

management, ou pas ?

En même temps, ce qu'elle propose, c'est quelque chose qui est un petit peu différent, c'est-à-dire qu'il y a notion de l'amont, mais c'est un amont, non pas pour étudier l'ensemble de la production mais pour choisir ce que l'on souhaite préserver, donc il y a cette notion de donnée essentielle, qui fait beaucoup débat dans la profession bien sûr, je ne vous le cache pas, qui est au cœur du rapport et qui revient complètement sur l'échantillonnage tel qu'on l'a pratiqué jusqu'à présent. Il y a en amont et en amont, on a toujours dit « il faut que l'archiviste soit en amont, associé, dès le premier moment, etc. », ça très bien, mais d'après elle ce n'est pas en amont pour faire du records management comme on l'a souvent entendu et notamment pour prévoir les échantillonnages, conserver, envisager... C'est plutôt être en amont pour dire ce qu'il veut, et puis sur le reste, je ne vais pas dire qu'on laisse courir, mais on est moins regardant.

Mais ça, ça fait un peu partie, finalement, des pratiques du records management, dans le sens où il y a les enjeux stratégiques, la gestion des risques, ce genre de choses, c'est inhérent au...

Oui, mais là elle rappelle que les services d'archives ont avant tout une mission sur les archives définitives. Donc en fait c'est aller en amont mais pour gagner du temps en aval, c'est plutôt ça, aller en amont pour dire « je veux ça, ça et ça » et après ça, une fois qu'on l'a, s'attacher à la description... En fait, il faut bien voir qu'un service d'archives, entre la collecte, l'évolution de la collecte d'un côté et la mise en ligne de l'autre, on a souvent minoré les moyens mis sur le classement, la conservation, on ne va pas dire qu'on ne l'a pas fait mais enfin souvent la collecte, parce qu'il y a une espèce d'impérieuse nécessité, les fonds arrivent, ils sont versés, il faut vider les placards, etc. La mise en ligne, parce qu'il y a une demande sociale, que ça répond aussi à des politiques des collectivités ou de l'État. Et puis entre les deux, des ressources qui sont du coup moins disponibles. Donc, son rapport a cet intérêt-là de rééquilibrer un peu les choses.

D'accord, j'ai juste une dernière petite question. Alors je me demandais si vous pouviez m'aiguiller sur les enjeux de l'exposition des données en ligne issues du traitement des documents d'archives, pour un secteur en particulier, celui de la recherche, pour les chercheurs, et notamment à travers le mouvement des humanités numériques. Comment est-ce que ces enjeux-là apparaissent ici ?

Alors, pour moi c'est quelque chose de très intéressant, j'avais participé à un groupe de travail sur Archives et territoires il y a quelques années, il y a un rapport qui doit se trouver quelque part sur le site du ministère de la Culture, où on avait essayé déjà, d'appréhender ça, on avait auditionné les chercheurs sur leur rapport aux ressources mises en ligne. Après, la question que j'ai, c'est finalement : est-ce que le public des chercheurs universitaires est si différent que ça dans ses usages que du grand public et est-ce que finalement il n'y a pas un peu cette tendance, bon là je schématise beaucoup, à considérer que tout ce qui n'est pas en ligne n'existe pas, par exemple. Souvent, ce qu'on perçoit chez le grand public, qui vous écrit en disant « je ne comprends pas, je ne trouve pas ceci, cela, comment se fait-il que ça ne soit pas en ligne ? » et effectivement vous êtes obligé de lui expliquer qu'il va falloir qu'il vienne en salle de lecture, tout n'est pas numérisé. D'ailleurs c'est une question qu'on nous pose tout le temps : « ah bon mais alors quand est-ce que vous aurez terminé de tout numériser ? ». Le public universitaire n'échappe pas forcément complètement à ça, même si on y trouve des gens qui sont férus d'actes originaux. Après, sur les humanités digitales, je pense que c'est encore très émergent dans les universités. On en parle beaucoup mais les moyens qui sont mis, voilà, moi je vois passer beaucoup de projets, chacun arrive avec sa base de données, son SIG, il faut numériser, bon... La pérennité de tout ça, la maintenance, l'évolution, l'actualisation, je pense que c'est des questions qui commencent à être prises en compte mais c'est très très émergent. En plus aujourd'hui on est sur un modèle de recherche qui est l'appel à projet, qui est une recherche avec un délai court, donc la problématique des données numériques c'est que c'est un temps plus long. Après, sur les usages, moi j'étais à une table ronde d'un groupe qui s'appelle Caillé, qui rassemblait des équipes de chercheurs qui ont mis en ligne des corpus

d'auteurs, donc c'est plutôt des littéraires, et qui travaillent ici avec une fille qui s'appelle Jessica de Bideran, quelqu'un de très intéressant sur les humanités digitales, elle est chargée de recherche et a pas mal travaillé sur cette question autour du patrimoine et là elle est sur le projet Mauriac en ligne, donc un projet d'humanités digitales de mise en ligne de tous les articles de presse de Pierre Mauriac. Après, ils faisaient quand même des constats sur qui utilise ces ressources mises en ligne, qui étaient assez mitigés. Donc il y avait un professeur de Grenoble, qui faisait partie d'une équipe qui a mis en ligne du Stendhal, alors il disait « on avait un pic de connexion tous les ans, la même semaine, depuis un site au fin fond des États-Unis » et en fait après analyse, c'était un prof, qui avait dû passer sur Grenoble pendant ses études et qui, tous les ans, faisait le même cours sur Stendhal, pendant une semaine, et il avait tous ses étudiants qui se connectaient en même temps au site et donc ils avaient un pic de consultation. Du coup, ce qui est intéressant, c'est que les humanités digitales amènent les universitaires à se poser la question de la réception de ce travail-là, parce que là pour le coup il y a quand même des sommes engagées qui sont visibles. Bon, un universitaire qui travaille chez lui, avec son ordinateur, ses fiches, il y a beaucoup de coûts masqués, mais par contre un universitaire qui veut mettre en ligne tel corpus etc, là les coûts ne sont plus du tout masqués, voilà. Autant sur les publications universitaires, il n'y avait pas forcément d'interrogation sur la réception, même si on sait que certains livres se vendent mieux que d'autres, là du coup sur ces corpus mis en ligne : quelle réception, quels usages, quels mésusages parfois ? Du coup, c'est des questions nouvelles pour beaucoup de chercheurs qui voyaient jusque-là l'objectif de la recherche, du groupe de recherche... finalement c'est la question de l'utilité secondaire, finalement ensuite tout ça, ça sert à qui ? Alors, il y a des gens qui font des choses intéressantes, qui essaient de mixer l'audiovisuel, de combiner toutes ces ressources du web pour ouvrir leur recherche, la diffuser, mais derrière, quelle est la fréquentation précise ? Par qui ? C'est des questions qui se poseront tôt ou tard... leur obsolescence *rires*.

Annexe 8 : les fonctionnalités du module référentiel du SAEM girondin

Fonctionnalités principales :

- Faciliter la gestion du SAE en permettant la réutilisation de données nécessaires aux différents processus d'archivage (versement, communication, restitution, etc.)
- Permettre la construction d'index de recherche interopérables à destination des gestionnaires du SAE et des communautés d'utilisateurs des productions électroniques au cours des différentes étapes de leur cycle de vie.
- Faciliter la cohabitation des exigences liées à la valeur probante des documents et à leur valeur informationnelle au sein d'un processus sécurisé.

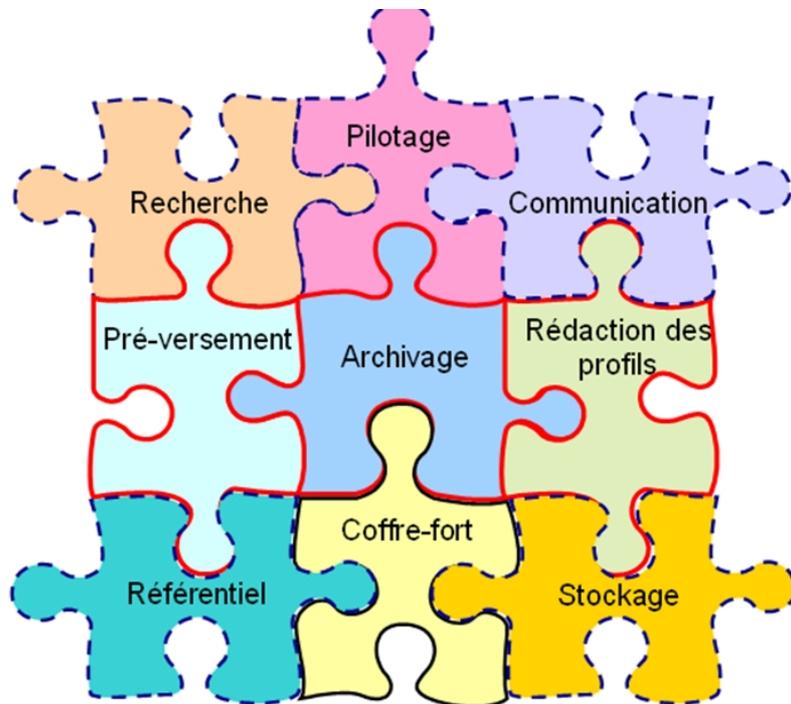
En matière de gestion des données de référence :

- L'import de thésaurus et des référentiels existants provenant des systèmes d'information, des collectivités ou du Web et leur maintien à jour en conservant et en exposant la provenance des données.
- La création de données de référence en leur attribuant des identifiants uniques et pérennes.
- L'import de descriptions normalisées de producteurs, notamment à travers le recours à des notices au format EAC-CPF rédigées dans le respect de la norme ISAAR-CPF.

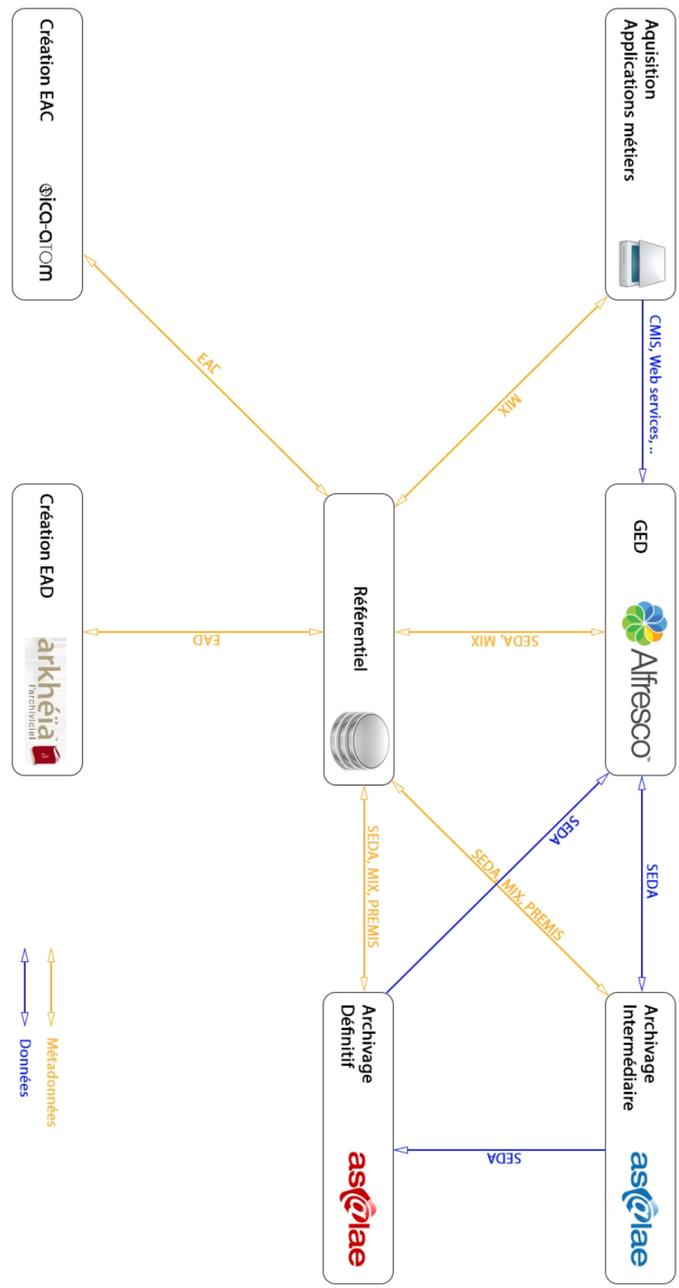
Pour ce qui est de la gestion des métadonnées :

- L'import d'une ontologie et sa mise à jour.
- L'import de métadonnées relatives au processus d'archivage.
- L'enrichissement de ces métadonnées à partir des données de référence.
- L'export des données dans un schéma standardisé (XML, RNG, etc.).

Annexe 9 : la vision modulaire propre au SAEM girondin



Annexe 10 : le positionnement du module référentiel dans l'architecture du SAEM

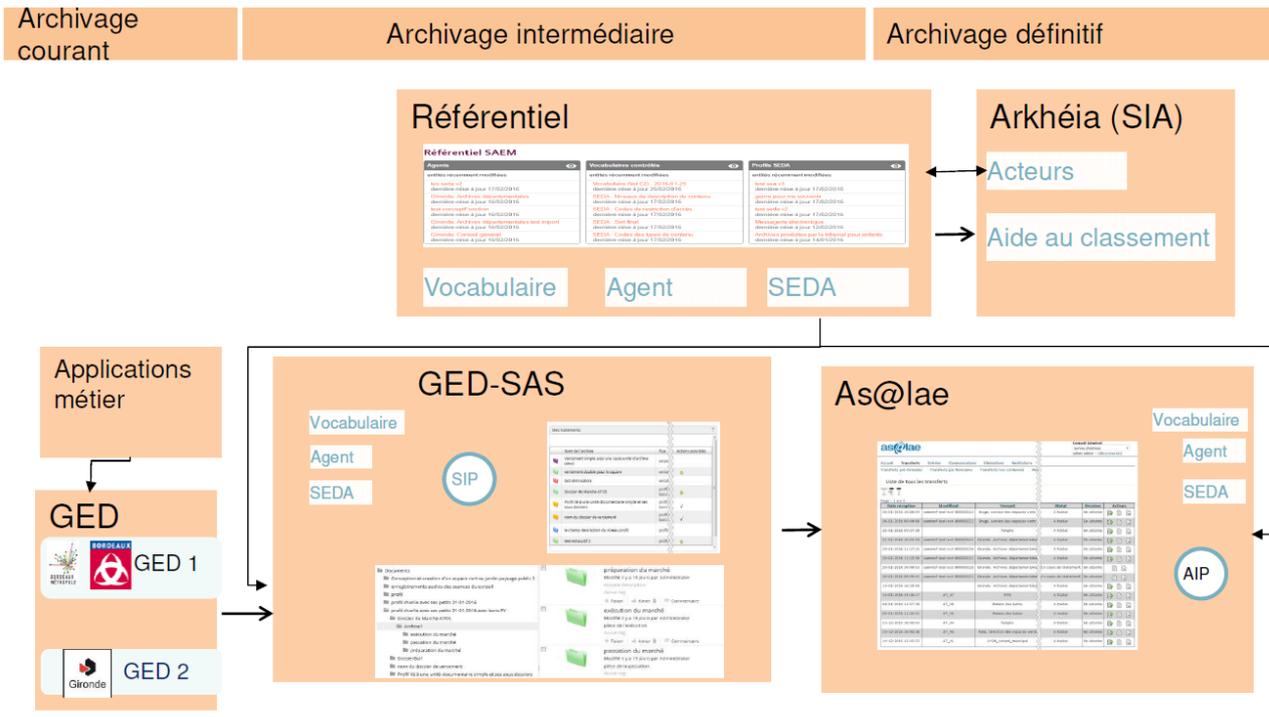


Annexe 11 : typologie des métadonnées du module référentiel

Typologie, famille, catégorie de métadonnées.	Schémas et standards
Métadonnées techniques	MIX / MPEG-7 / TextMD
Métadonnées administratives (de gestion)	PREMIS / SEDA / XML-EAC / XML-EAD/ PROV-O
Métadonnées descriptives	XML-EAD / XML-EAC DUBLIN CORE / SKOS
Métadonnées de structure	METS / RDF / MPEG21 et MPEG21-DIDL / XFDU / OWL

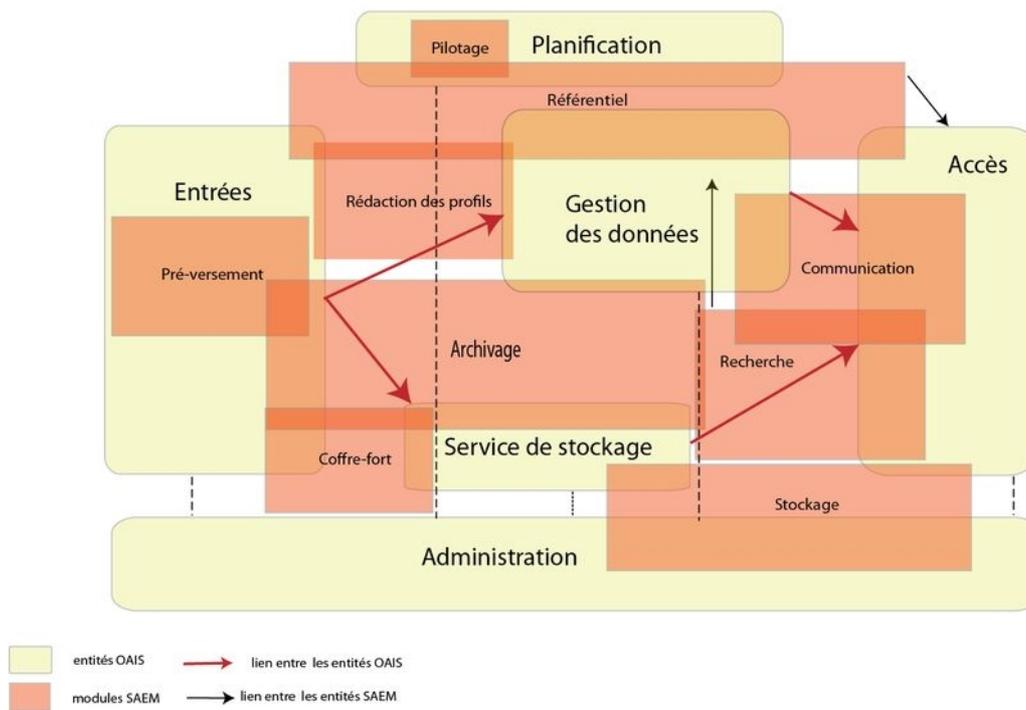
Typologie des métadonnées du référentiel.

Annexe 12 : les modules du SAEM girondin et le cycle de vie des archives



Annexe 13 : articulation des modules du SAEM avec le modèle OAIS

Recouvrement des entités
modèle conceptuel OAIS - briques fonctionnelles issues de la vision modulaire

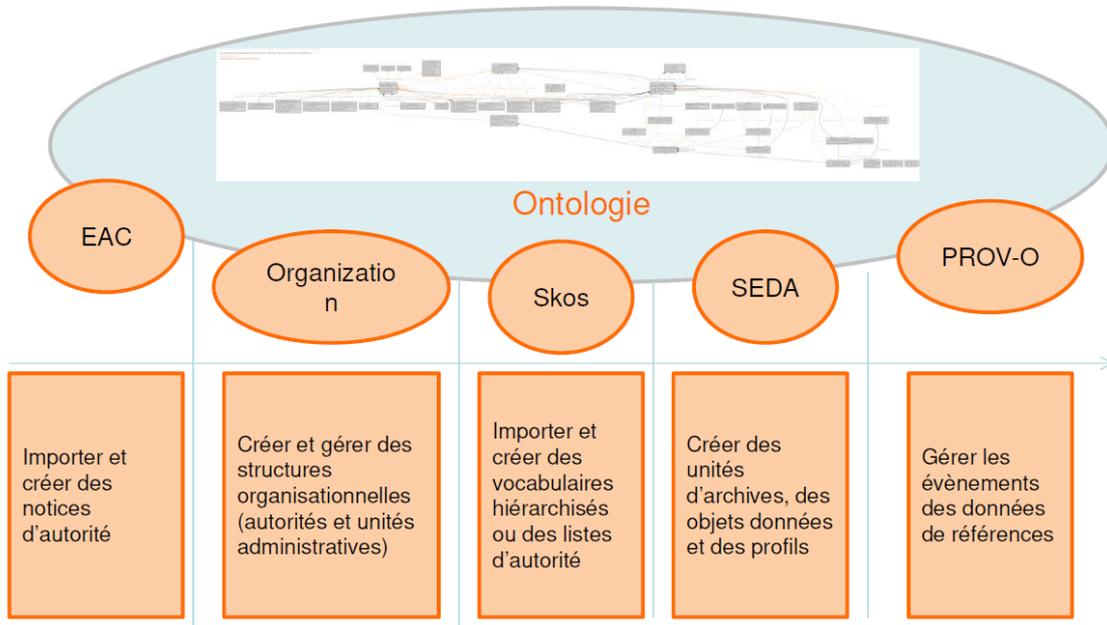


Annexe 14 : répartition des rôles pour la mise en production du SAEM girondin

Tableau de répartition des rôles pour la mise en production du SAEM (applications et données au Département, réplication à Bordeaux Métropole)

Collectivité productrice et fonds nativement numériques concernés	Autorité d'archivage	Autorité de contrôle (contrôle scientifique et technique de l'Etat)	Opérateur d'archivage chargé de l'exploitation	Opérateur d'archivage chargé de la réplication des données	Opérateur de versement
Département de la Gironde : production des services départementaux et organismes publics rattachés	Archives départementales de la Gironde	Directrice des Archives départementales de la Gironde	Département de la Gironde (direction des systèmes d'information)	Bordeaux Métropole (direction générale du numérique et des systèmes d'information)	Département de la Gironde (direction des systèmes d'information)
Services déconcentrés de l'Etat (cas de figure déjà existant : Cadastre, boîtes mail)	Archives départementales de la Gironde	Directrice des Archives départementales de la Gironde	Département de la Gironde (direction des systèmes d'information)	Bordeaux Métropole (direction générale du numérique et des systèmes d'information)	Direction des systèmes d'information du service ou tiers
Communes ou tout autre personne chargée d'une missions de service public du département de la Gironde : production des services et établissements rattachés	Archives départementales de la Gironde	Directrice des Archives départementales de la Gironde	Département de la Gironde (direction des systèmes d'information)	Bordeaux Métropole (direction générale du numérique et des systèmes d'information)	Direction des systèmes d'information du service ou tiers
Autre département : production des services départementaux et organismes publics rattachés	Archives départementales concernées	Directeur des Archives départementales concernées	Département de la Gironde (direction des systèmes d'information)	Bordeaux Métropole (direction générale du numérique et des systèmes d'information)	Département concerné (direction des systèmes d'information)
Région : production des services régionaux et établissements rattachés	Archives régionales	Directeur des Archives départementales concernées	Département de la Gironde (direction des systèmes d'information)	Bordeaux Métropole (direction générale du numérique et des systèmes d'information)	Région concernée (direction des systèmes d'information)

Annexe 15 : l'articulation des différentes couches du modèle de données du SAEM



Annexe 16 : les user stories issues du backlog de Logilab

1. « En tant qu'archiviste je veux pouvoir renseigner le fournisseur et/ou le producteur dans le module entrées à partir du module référentiel afin d'éviter une double saisie et utiliser mes données de référence. »
2. « En tant qu'archiviste je veux pouvoir récupérer une liste d'acteurs de type services producteurs dans le module d'aide au classement pour éviter d'avoir à ressaisir la description de cet acteur »
3. « En tant qu'archiviste je veux pouvoir récupérer une liste de vocabulaires dans le module d'aide au classement pour indexer mes archives à partir des concepts disponibles. »
4. « En tant qu'utilisateur du module référentiel je veux pouvoir attribuer des identifiants pérennes aux objets gérés par le référentiel afin de pouvoir suivre leur cycle de vie dans le cadre de leur utilisation. »
5. « En tant qu'archiviste je veux qu'un identifiant soit attribué aux unités d'archives de manière automatique afin de pouvoir importer une unité d'archives sans avoir forcément à changer son titre. »
6. « En tant qu'administrateur fonctionnel je veux pouvoir importer des thésaurus et des listes d'autorités existant au format SKOS/XML afin de centraliser les données de référence utilisées dans le cadre du SAEM. »
7. « En tant qu'administrateur fonctionnel du SAEM je veux pouvoir créer des listes d'autorités ou des thésaurus afin de les utiliser dans le cadre du SAEM ou du SI. »
8. « En tant qu'administrateur fonctionnel je veux pouvoir exporter les vocabulaires contrôlés gérés dans le référentiel afin de les rendre disponibles pour la consommation d'autres modules du SAEM ou du SI. »
9. Mettre à jour les sources de vocabulaires des informations d'agent lors d'import de vocabulaires

« Lorsque j'importe un agent, certaines informations utilisent une source de

vocabulaire (par exemple, les fonctions). Lorsque le vocabulaire contrôlé n'est pas présent dans le référentiel, on stocke l'URI (externe) du concept. Lorsque j'importe *a posteriori* le vocabulaire référencé dans une information d'un agent, je voudrais que sa source de vocabulaire pointe vers l'entité "concept" du référentiel issue de l'import et non plus vers une URI externe. »

10. Lier un vocabulaire issu d'une liste d'autorités ou d'un thésaurus à une fiche acteur

« En tant qu'archiviste je souhaite pouvoir lier un terme d'indexation issu d'une liste d'autorités ou d'un thésaurus présent dans le référentiel à une notice d'autorité (acteur) afin d'enrichir cette notice d'autorité (acteur) »

11. Export SKOS XML

« On veut pouvoir exporter au format SKOS XML les thésaurus et listes d'autorité présentes dans cubicweb. En tant qu'archiviste je veux pouvoir exporter tout vocabulaire créée ou importer dans le référentiel au format SKOS pour permettre l'échange d'information entre les différentes briques logicielles du SAE (format d'échange SKOS pour les vocabulaires) »

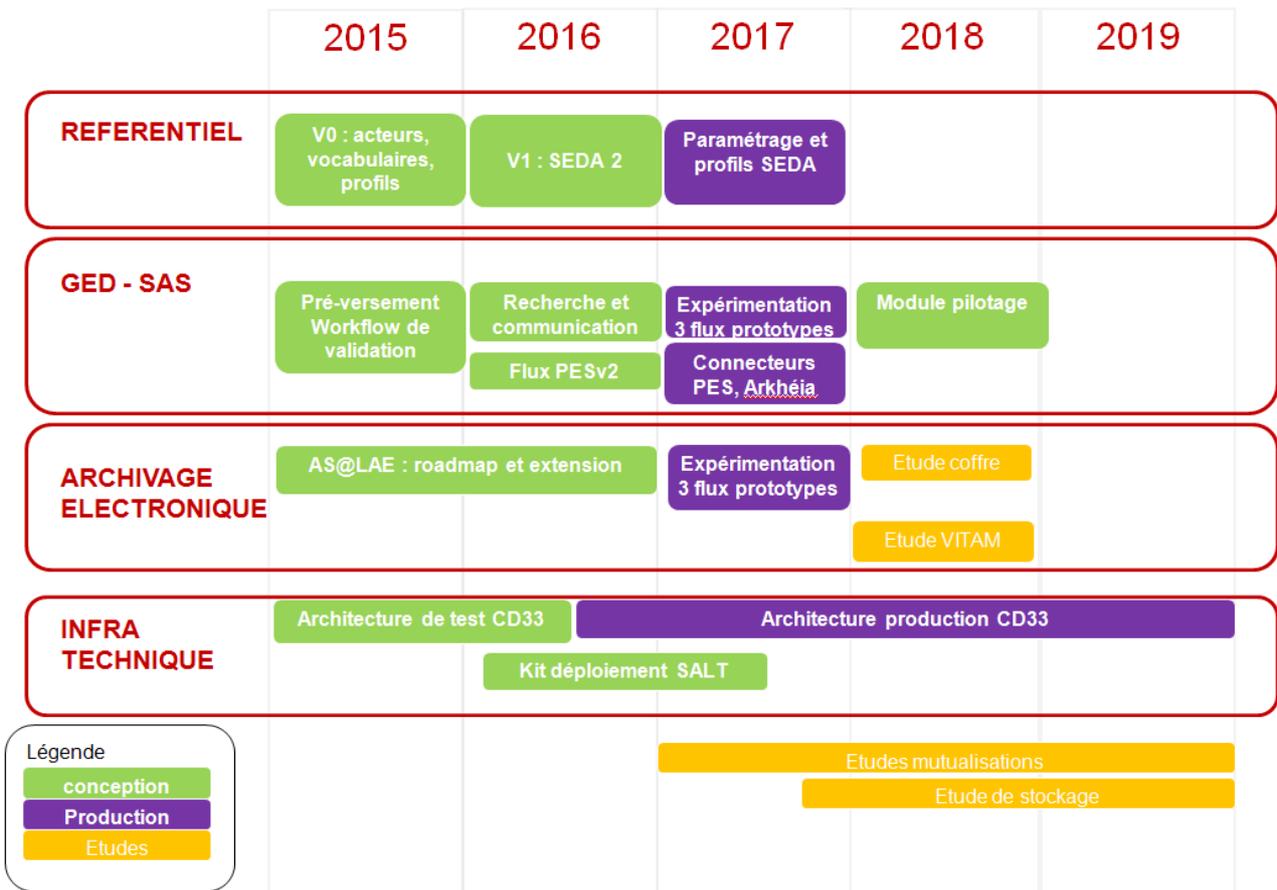
12. Utiliser l'entité prov : activity pour tracer les événements du module référentiel :

« En tant qu'architecte du module référentiel je veux que tous les événements du module référentiel soient structurés en tant qu'entité de type prov : activity afin de pouvoir les interroger. »

13. Exposition modèle de données prov-o dans les vues RDF

« En tant qu'utilisateur je veux que les entités du modèle de données Prov-o soient visibles dans les vues d'export RDF fournies par le référentiel Pour ce qui concerne les agents (notice d'autorité) on a convenu de mapper les maintenanceEvent en import vers Prov-o mais ensuite toutes les actions effectuées sur les notices d'autorité doivent faire l'objet d'une transformation en prov-o dans les vues RDF. Pour les vocabulaires, on aimerait pouvoir conserver l'historique des créations et des modifications des vocabulaires en prov-o si ce modèle est déjà implémenté. Sinon a minima on souhaite que soient exposées en SKOS (DCT) les événements de création et de modification sur les concepts. »

Annexe 17 : le planning du projet SAEM



Annexe 18 : la vision fonctionnelle cible du SAEM

