

CIBAHN' 2016

Archivage numérique et réseaux

Actes de la première édition du colloque
international sur les Bibliothèques et Archives à
l'ère des Humanités Numériques

Tunis, 19 et 20 octobre 2016

Sous la direction de
Abderrazak Mkadmi
Imad Saleh
Khaled Habchi

Editions de l'institut supérieur de documentation
2016

Archivage numérique et réseaux

Actes du colloque international sur les bibliothèques et archives à l'ère des humanités numériques/ Sous la direction de Abderrazak Mkadmi, Imad Saleh et Khaled Habchi. Tunis : Institut supérieur de documentation, 2016. – 250 pages

ISBN : 978-9973-904-13-3

Edition Octobre 2016

Institut supérieur de documetation ; ISD

Université de la Manouba. Tunis

Adresse : Campus universitaire de la Manouba ; 2010 Tunisie

Tél. 216. 71. 601. 050

Fax 216. 71. 600. 200

Mail isd@isd.rnu.tn

Site : www.isd.rnu.tn

ISBN : 978-9973-904-13-3

©2016, ISD

Tous droits, même de reproduction d'extraits, de reproduction photomécanique ou de traduction réservés.

COMITES DU COLLOQUE

COMITÉ SCIENTIFIQUE

Ghislaine Azémard, Université Paris8, France

Bruno Bachimont, Université de Technologie de Compiègne,
France

Claude Baltz, Université Paris 8, France

Mohamed Ben Romdhane, ISD, Université de la Manouba,
Tunisie

Azzedine Bouderbane, Institut de Bibliothéconomie, Université
Constantine 2, Algérie

Besma Bsir, ISD, Université de la Manouba, Tunisie

Nasreddine Bouhai, Université Paris8, France

Aïda Chebbi, ISD, Université Manouba, Tunisie

Abderrazak Dourari ; Université Alger2, CNPLET, Algérie

Jacques Ducloy, DRRT Lorraine, France

Raja Fenniche, ISD, Université de la Manouba, Tunisie

Wahid Gdoura, ISD, Université de la Manouba, Tunisie

Khaled Habchi, ISD, Université de la Manouba, Tunisie

Madjid Ihadjadene, Université Paris8, France

Hédi Jallab, Archives nationales de Tunisie, Tunisie

Najoua Jrad, ISD, Université de la Manouba, Tunisie

Abderrazak Mkadmi, ISD, Université de la Manouba, Tunisie

Tarek Ouerfelli, ISD, Université Manouba, Tunisie

Everardo Reyes, Université de Paris8, France

Imad Saleh, Université Paris8, France

Yousra Séghir, ISD, Université de la Manouba, Tunisie

Peter Stockinger, INALCO, France

Rachid Zghibi, ISD, Université de la Manouba, Tunisie

Khansa Zghidi, ISD, Université de la Manouba, Tunisie

Khaldoun Zreik, Université Paris8, France

COMITÉ D'ORGANISATION LOCAL

Mohamed Ben Romdhane, ISD, Université de la Manouba, Tunisie

Besma Bsir, ISD, Université Manouba, Tunisie

Kais Chihi, TechnoPark Manouba

Samira Ghourabi, ISD, Université Manouba, Tunisie

Khaled Habchi, ISD, Université de la Manouba, Tunisie

Emna Madani, ISD, Université de la Manouba, Tunisie

Abderrazak Mkadmi, ISD, Université de la Manouba, Tunisie

Yousra Séghir, ISD, Université Manouba, Tunisie

Table des matières

Introduction	7
Abderrazak Mkadmi.....	7
Imad Saleh.....	7
Khaled Habchi.....	7
Modélisation des collections numériques : un projet de recherche-création sur l’Odyssée	10
Arnaud Laborderie	10
Ecosystèmes de connaissances pour les archives numériques ..	26
Samuel Szoniecky	26
Le rôle de records management dans la gestion stratégique de l’information des organisations.....	38
Anna Lezon Rivière	38
Services de gestion de données de recherche en Suisse : Vers une collaboration élargie & une offre exhaustive	56
Pierre-Yves Burgi.....	56
Eliane Blumer.....	56
Basma Makhoulf-Shabou.....	56
Lydie Echernier	56
Pratiques et enjeux de la documentation numérique personnelle	84
Nasreddine Bouhai	84
Is Digital Humanities changing the learner-instructor interaction: an inquiry of social media use in archival education in Algeria?.....	104
Behdja Boumarafi	104
La médiation numérique et l’avenir des métiers de l’information-documentation face à l’inflation du document numérique.....	116
Mor DIEYE.....	116
Humanités numériques et métier d’archiviste : quelle articulation possible ?.....	133

Mohamed Lat Diop.....	133
Les Archives à l'ère des Humanités Numériques : Quelles mutations ?	154
Fatma Bilel & Abderrazak Mkadmi.....	154
Bibliothèques numériques et crowdsourcing : analyses bibliométriques et text mining.....	176
Mathieu Andro (1,2), Imad Saleh (2), Samuel Szoniecky (2)	176
Archives et transition démocratique à l'ère de des humanités numériques : la situation en Tunisie post-révolutionnaire	186
Yousra Seghir, Khaled Habchi.....	186
An Interdisciplinary Approach for Cultural Heritage Valorisation and Visualization.....	209
Basma Makhoulouf Shabou ^a	209
Maria Sokhn ^b	209
Alex Olivieri ^b	209
Expériences de numérisation à la Télévision Tunisienne : processus et défis Etude de cas : les supports type 02 pouces	235
Kaouthar HAMED	235

Introduction

Abderrazak Mkadmi (1)

Imad Saleh (2)

Khaled Habchi (3)

(1, 2) Unité de recherche BNP, ISD, Université de la Manouba
(Tunisie)

amkadmi@gmail.com, khaledhabchi@gmail.com

(3) Laboratoire Paragraphe, Université Paris 8 (France)

idsaleh@orange.fr

Pour sa première édition, le colloque CIBAHN (Bibliothèques et Archives à l'ère des humanités numériques) se tiendra du 19 au 20 octobre 2016 sous le thème « *Archivage numérique et réseaux* »

Réunissant des chercheurs de tous horizons, la première édition de ce colloque se veut être un espace d'échange et de valorisation des recherches dans le domaine de la gestion des archives, plus précisément celles qui portent sur les transformations du métier d'archiviste à l'ère du Web social.

Dans cette nouvelle réalité virtuelle, les contenus de la sphère publique sont entremêlés avec ceux de la sphère privée de sorte qu'il est de plus en plus difficile de définir avec certitude le statut des objets informationnels véhiculés par le Web. En outre, le Web social regorge de nouveaux genres documentaires devant être étudiés dans une perspective interdisciplinaire voire même transdisciplinaire dans l'objectif de développer de nouveaux modèles d'acquisition, de traitement, de conservation et de mise en valeur propres aux contenus numériques de nature dynamique et protéiforme.

Les réseaux sociaux séduisent les milieux organisationnels. Des blogs corporatifs, des pages Facebook, des comptes Twitter, etc. deviennent la nouvelle vitrine des organisations sur le Web. Initialement les réseaux sociaux sont utilisés comme outils de

marketing, ils bouleversent aujourd'hui le rapport à l'information et créent un nouveau modèle de gestion du savoir et du savoir-faire dans les organisations. Cette « technologisation » est appuyée par une société de « requête » qui donne de l'importance à l'accès instantané et néglige souvent les problématiques liées à la préservation et la réutilisation à long terme des contenus numériques. Ceux-ci sont notamment le produit des activités organisationnelles sur les réseaux, avec tout ce que cela implique comme questions connexes liées au droit, à l'éthique, à la protection de la vie privée, à la sécurité et à la fiabilité des traces numériques devant être conservées.

La conservation des traces des activités qu'elles soient personnelles ou organisationnelles menées sur le Web social ou via ses outils représente un enjeu de taille, et ce, à plusieurs niveaux. Au niveau technique, archiver de nouveaux genres documentaires dont la nature est hybride et évolutive exige de nouvelles compétences. Au niveau juridique et éthique, traiter des documents Web comportant souvent des données privées nécessite de nouvelles approches et méthodes de capture, de traitement et de réutilisation.

Nous proposons dans cette première édition du colloque une réflexion générale sur toutes ces questions liées aux avancées technologiques de diffusion, de mise en valeur et de pérennisation des contenus à l'ère des humanités numériques et de leur impact sur le métier de l'archiviste tant au niveau professionnel, que culturel et social.

Modélisation des collections numériques : un projet de recherche-crédation sur l'Odysée

Arnaud Laborderie

Bibliothèque nationale de France, chaire Unesco ITEN,
Université Paris 8, Laboratoire Paragraphe

arnaud.laborderie@gmail.com

Résumé

Le numérique redéfinit les espaces du livre et de la bibliothèque. À la double nature du livre imprimé — œuvre et objet (Chartier, 2009) — répondent les trois dimensions du livre numérique — discursive, computationnelle et réticulaire (Mounier, 2010). Évoquant la double nature du livre, Michel Foucault (1969) parle d'un « nœud dans un réseau » car l'œuvre est prise dans un système de renvois qui repousse les frontières de l'objet : réseau de références implicites ou explicites, réseau intertextuel fait, selon Genette (1980), d'hypertextes (réécritures) et de métatextes (commentaires), réseau de lecteurs pour qui l'œuvre est « ouverte » (Eco, 1962) à toutes les relectures et interprétations. Aussi peut-on attendre des bibliothèques numériques qu'elles donnent non seulement accès aux œuvres, mais aussi qu'elles restituent leurs formes et actualisent leur intertextualité. Deux approches nous paraissent répondre à cette attente : une approche orientée « données », celle de data.bnf et des humanités numériques ; une approche orientée « utilisateurs », celle de la médiation numérique et de la valorisation culturelle. À côté des approches statistiques portées par les humanités numériques, nous souhaitons exposer une approche tournée vers l'« expérience utilisateur ». C'est l'enjeu d'un programme de recherche de la BnF consacré à la modélisation des collections numériques : expérimenter, avec une équipe de chercheurs et des étudiants de master, de nouveaux modes de médiations aux ressources numériques, dans le cadre d'un atelier-laboratoire IDEFI-CréaTic. Nous adoptons

ainsi une posture pragmatique à travers laquelle articuler recherche et création. Pour cela, l'*Odyssee* apparaît comme une matière exemplaire. Première « œuvre hypertextuelle » selon Genette, elle se présente comme l'archétype d'une « œuvre ouverte ». De l'oral à l'écrit, l'œuvre, remédiatisée de nombreuses fois au cours des siècles, transcende ses différents supports. Comment remédier à l'*Odyssee* à travers les collections patrimoniales de la BnF ? Peut-on saisir les différentes dimensions de l'œuvre dans une interface innovante ? Il ne s'agit pas de concevoir une édition augmentée qui serait exhaustive, totale, mais de restituer la richesse de l'œuvre et sa complexité à travers une « expérience utilisateur » constituant une médiation au corpus odysseén.

Nous proposons une communication au colloque CIBAHN en deux temps : un premier temps théorique, avec l'exposé de la problématique du livre augmenté et de la bibliothèque numérique, des approches « données » et « utilisateurs », du projet de recherche BnF et de la méthodologie de l'atelier-laboratoire, suivi d'un second temps expérimental avec la démonstration du prototype.

Mots-clés : bibliothèque numérique, collections numérisées, valorisation patrimoniale, médiation numérique et culturelle, *Odyssee*, livre augmenté, édition numérique.

Notice biographique

Arnaud Laborderie est chef de projet au département des Éditions de la Bibliothèque nationale de France. Actuellement, il est chargé d'un projet d'éditorialisation de la bibliothèque numérique Gallica à destination des publics scolaires et des enseignants, les « Essentiels de la littérature » : <http://gallica.bnf.fr/essentiels/>

Chercheur à la chaire ITEN (Innovation, Transmission, Édition Numériques) de l'Unesco et enseignant dans le master CEN (Création et Édition Numériques) de l'université Paris 8, il expérimente les nouvelles formes et nouveaux usages du livre dans le cadre d'un atelier-laboratoire IDEFI-CréaTic : <http://idefi-creatic.net/>

1/ Problématique et concept du « livre augmenté »

Le numérique redéfinit les espaces du livre et de la bibliothèque. Qu'est-ce qu'un livre à l'ère numérique ? Que peut-on attendre des bibliothèques virtuelles ? Quelles médiations nouvelles la numérisation permet-elle aux corpus et aux œuvres ? Ces questions appellent pour nous deux réponses : une réponse théorique par la définition du concept de « livre augmenté » ; une réponse pragmatique par la conception d'un prototype dans le cadre d'un atelier-laboratoire.

Le livre de l'ère numérique sera-t-il un « livre augmenté » à l'image de cet « homme augmenté » que nous promettent, pour le meilleur et pour le pire, les prophètes des nouvelles technologies ?

Augmenter l'expérience de lecture et l'accès aux œuvres

Au-delà des applications émergentes qui usent de la réalité augmentée pour associer le livre papier avec des extensions virtuelles sur des téléphones mobiles ou des tablettes numériques, nous proposons une définition du « livre augmenté » comme « augmentation de l'expérience de lecture » : le texte est augmenté par les ressources du dispositif, lesquelles peuvent combiner plusieurs médias. Ainsi l'œuvre augmentée n'est-elle pas, selon nous, le propre du numérique : elle s'inscrit dans la continuité des technologies intellectuelles qui ont perfectionné le livre depuis les listes et tableaux sumériens (Goody, 1979). De ce point de vue, les évolutions techniques de la page ont produit des livres augmentés dans divers supports et à différentes époques, parmi lesquels citons, à la suite d'Anthony Grafton (2012), la *Chronique* d'Eusèbe de Césarée (325), la *Chronique de Nuremberg* (1493) ou le Dictionnaire de Bayle (1697). L'expérience de lecture s'y trouve en effet augmentée par des innovations techniques et des enrichissements de contenu. Si, dans l'histoire de l'édition fleurissent les œuvres « augmentées » par des ajouts d'images ou des appareils critiques, l'augmentation, dans notre approche, conjugue enrichissements et innovations pour renouveler le mode de lecture et l'accès au texte.

Que pourrait être alors une édition numérique augmentée ? Apportons des éléments de réponse théoriques, en mettant en relation la double nature du livre imprimé — œuvre et objet

(Chartier, 2009) — avec les trois dimensions du livre numérique — discursive, computationnelle et réticulaire (Mounier, 2010).

C'est en s'inscrivant dans la tradition des Lumières portée par Emmanuel Kant ou Denis Diderot¹ que Roger Chartier affirme la double nature matérielle et intellectuelle du livre imprimé. Dans son cours au collège de France, l'historien parle poétiquement du « corps » et de l'« âme » du livre qui est à la fois un objet, incarné dans une forme matérielle (celle du *codex*, par exemple), et une œuvre (discours, récit ou pensée), c'est-à-dire une forme immatérielle que nous pourrions caractériser d'« intellectuelle » ou de « spirituelle ».

L'œuvre, âme du livre, habite plusieurs corps : ce sont les formes qu'elle a visitées, chaque nouvelle forme se souvenant, en quelque sorte, de la précédente et la dépassant, selon la théorie de la « remédiatisation » (Bolter et Grusin, 1999), pour laquelle les formes médiatiques se transforment les unes dans les autres et gardent trace, implicite ou explicite, de ses transformations : un changement de support, une nouvelle interface, une redistribution des contenus avec du commentaire, de l'illustration, etc. Ainsi l'œuvre s'inscrit-elle dans le *continuum* de ses remédiatisations, lesquelles produisent des objets porteurs de sens dans leur matérialité même et dans leurs composantes (Chartier, 2012).

L'œuvre, bien que close dans sa forme matérielle, enclose dans l'objet livre, reste une « œuvre ouverte », ouverte, nous dit Umberto Eco, aux multiples lectures et interprétations : « Toute œuvre d'art alors même qu'elle est une forme achevée et close dans sa perfection d'organisme exactement calibré, est ouverte au moins en ce qu'elle peut être interprétée de différentes façons, sans que son irréductible singularité soit altérée. Jouir d'une œuvre d'art revient à en donner une interprétation, une exécution, à la faire revivre dans une perspective originale. » (Eco, 1962)

Pour Gérard Genette (1980), les œuvres littéraires tissent entre elles des liens intertextuels qui peuvent être des réécritures (hypertextes) ou des commentaires (métatextes). On voit ainsi

¹ Diderot, *Lettre sur le commerce de la librairie* (1763) sur Gallica : <http://gallica.bnf.fr/ark:/12148/bpt6k6443411c> ; Kant, *Principes métaphysiques de la doctrine du droit* (1797) sur Gallica : <http://gallica.bnf.fr/ark:/12148/bpt6k6571553h>. « Qu'est-ce qu'un livre ? » <http://gallica.bnf.fr/ark:/12148/bpt6k6571553h/f65>

que l'œuvre s'étoile en quelque sorte dans ses formes matérielles et intellectuelles : étoilement du corps, par ses multiples formes, étoilement de l'âme, par ses intertextes et ses références.

Évoquant cette double nature du livre, en 1969 dans *L'archéologie du savoir*, Michel Foucault souligne que « les marges d'un livre ne sont jamais nettes ni rigoureusement tranchées : par-delà le titre, les premières lignes et le point final, par-delà sa configuration interne et la forme qui l'autonomise, il est pris dans un système de renvois à d'autres livres, d'autres textes, d'autres phrases : nœud dans un réseau. » (Foucault, 1969 : 35)

Cette vision du livre comme un « nœud dans un réseau » évoque le concept de « rhizome » de Deleuze et Guattari (*Mille plateaux*, 1980) : « Le système radicelle, ou racine fasciculée, est la seconde figure du livre, dont notre modernité se réclame volontiers. Cette fois, la racine principale a avorté, ou se détruit vers son extrémité ; vient se greffer sur elle une multiplicité immédiate et quelconque de racines secondaires qui prennent un grand développement. » (Deleuze et Guattari, 1980 : 12) De l'arbre-racine (première figure du livre) au rhizome-canal (seconde figure), les éléments deviennent des multiplicités : ils forment des plateaux qui, articulés les uns aux autres, produisent des combinaisons. Ainsi le rhizome constitue-t-il pour nous une représentation mentale du livre « nœud dans un réseau » et « œuvre ouverte ». Il s'impose comme métaphore et représentation du livre augmenté dont les plateaux interconnectent deux plans : le plan matériel des remédiatisations et le plan intellectuel de l'intertextualité, des références et des interprétations. L'œuvre augmentée peut-elle combiner ainsi, autour du texte, une lecture en plateau de ses formes et de ses relations sémantiques ?

Cette double nature du livre imprimé, le livre numérique le transcende et dépasse lui aussi, en intégrant un système d'information, une base de données ou une bibliothèque numérique. Paradoxalement, la numérisation restitue matérialité et immatérialité en dématérialisant : elle donne à voir les formes, grâce à la visualisation en mode image — qui expose le corps du livre — et en mode texte qui permet d'en saisir, d'en approcher l'âme : c'est ce que propose Gallica, la bibliothèque numérique de la BnF et de ses partenaires.

La double nature du livre imprimé se révèle ainsi à travers les trois dimensions du livre numérique : discours, programme et réseau. Pour Pierre Mounier en effet on peut « considérer l'environnement numérique de trois dimensions complémentaires et essentielles, qui en constituent la réalité même : la dimension discursive – le discours, la dimension computationnelle – le programme, et la dimension réticulaire – le réseau social. Il est dès lors possible d'examiner non seulement le livre numérique, mais finalement tout objet numérique sous l'angle de chacune de ces trois dimensions. Tout objet numérique possède trois faces : il est un programme, il est un discours, et il est aussi un élément du réseau. » (Mounier, 2010).

Nous avons vu que l'œuvre était elle-même discours et réseau : c'est donc la dimension computationnelle qui caractérise le livre numérique et permet d'amplifier les dimensions discursive et réticulaire. L'œuvre augmentée, appréhendée comme un livre-réseau, dépasse alors ses formes numérisées pour devenir édition électronique qui, rappelons-le, « repose sur cinq piliers : la structuration de l'information, la documentation de l'information, l'optimisation des conditions de lecture, l'appropriation par le lecteur et le développement des interopérabilités. » (Dacos et Mounier, 2010 : 108). Sur le web, ces pratiques sont désignées sous le terme d'« éditorialisation », lequel « met l'accent sur les dispositifs technologiques qui déterminent le contexte d'un contenu et son accessibilité. » (Vitalo-Rossati et Sinatra, 2014 : 9).

En articulant ainsi discours, programme et réseau, le numérique permet de concevoir des livres augmentant l'expérience de lecture et la réception de l'œuvre. (Laborderie, 2015).

Le livre augmenté et la bibliothèque numérique

Les trois dimensions discursive, computationnelle et réticulaire redéfinissent la place du livre dans la bibliothèque. Engagées en premier lieu dans des politiques de numérisation avec rétroconversion des catalogues, les bibliothèques numériques ont permis l'accès des œuvres à distance sous forme numérique. Avec des millions de ressources en ligne, l'enjeu réside désormais dans l'enrichissement des métadonnées et l'interconnexion des catalogues et des documents : c'est ce que propose data.bnf.fr en s'appuyant sur les notices d'autorité dans une perspective potentiellement exhaustive avec près de deux

millions d'auteurs indexés et quelque 200 000 œuvres référencées. Ainsi manuscrits, éditions, illustrations ou spectacles sont-ils associés pour une même œuvre. Aujourd'hui, la problématique concerne l'accroissement exponentiel des ressources et la difficulté à circuler dans des corpus complexes.

Dès lors que peut-on attendre de la bibliothèque virtuelle ? Qu'elle donne non seulement accès aux œuvres, mais aussi qu'elle restitue leurs formes et actualise leur intertextualité. Deux approches nous paraissent répondre à cette attente : une approche orientée « données », celle de data.bnf et des humanités numériques ; une approche orientée « utilisateurs », celle de la médiation numérique et de la valorisation culturelle. Ces deux orientations relèvent de logiques et d'expertises différentes pour des publics différenciés.

2/ Les approches « données » et « utilisateurs » : enjeux des humanités numériques et programme de recherche

L'approche orientée « données » repose d'une part sur l'indexation et les métadonnées, avec des outils comme data.bnf.fr², et d'autre part sur l'encodage d'entités nommées par un balisage de type TEI (Text Encoding Initiative). Œuvres et documents sont ainsi considérés comme des données, lesquelles permettent une visualisation des corpus sous forme d'arbres, de cartes ou de graphes (Moretti, 2008). Potentiellement exhaustive, l'approche « données » propose un mode organisation des ressources, grâce à des ontologies, avec la difficulté de hiérarchiser et d'attribuer une valeur qualitative aux données ainsi traitées en masse : l'objectif est d'organiser la complexité et d'en faire des représentations pour interpréter.

Les humanités numériques appliquent ainsi au livre la statistique et la visualisation de données. En passant du *close reading* au *distant reading* (Moretti, 2008), elles mettent en perspectives, offrent des visions de masse et des approches diachroniques qui renouvellent la lecture et l'interprétation des textes. Par des fouilles lexicales de textes, l'algorithme peut modéliser, à l'intérieur même d'une œuvre, les relations entre les personnages, les déplacements dans le récit, les occurrences des

² Le projet data.bnf.fr propose des pages de référence sur les auteurs, œuvres et thèmes qui regroupent les informations de Gallica et celles issues des différents catalogues bibliographiques de la BnF. En ligne : <http://data.bnf.fr/>

mots, les figures de style, etc. Le lecteur appréhende ainsi les textes par des diagrammes, des représentations. Les humanités numériques changent le regard sur les œuvres et sur les corpus, considérés à une autre échelle. Elles renouvellent l'approche des textes, revisitent les sciences humaines par des pratiques pluridisciplinaires, régénèrent la philologie et l'herméneutique.

Ces formes de médiations, assistées par l'algorithme, s'affirment comme des pratiques savantes, à la fois de traitement des corpus et d'interprétation, partagées par une communauté scientifique. « Faire des humanités numériques » reste en effet une pratique de recherche (Berra, 2012).

Pour une approche orientée « utilisateurs »

Une autre approche plus éditoriale, tournée vers la médiation « grand public » et la valorisation culturelle, s'attache à « l'expérience utilisateur » et au design des interfaces. « L'expérience utilisateur correspond à l'expérience vécue par une personne lorsqu'elle utilise un produit, un service, un outil. Dans le contexte des métiers interactifs, l'outil en question est le site Internet, l'internaute va donc vivre une expérience lors de l'utilisation du site. La qualité du site Internet va déterminer la qualité de l'expérience utilisateur. » (Pignier et Drouillat, 2004 : 199)

Appréhender le livre de cette manière c'est lui appliquer les principes de la conception d'interface, « orientée utilisateurs ». Partir de l'usage et considérer l'utilisabilité (Baccino, 2004) pour concevoir le livre-interface : le lecteur, ce n'est plus l'algorithme mais l'utilisateur. Aussi importe-t-il de considérer la page comme une interface et l'interface comme une page : le « parcours utilisateur » devient alors un mode de lecture. Cette approche organise une sélection de ressources dans des parcours sémantiques où l'« expérience utilisateur » constitue une médiation au corpus.

Peut-on associer les orientations « utilisateurs » et « données » en proposant une interface au travail d'encodage et d'interconnexion des données ? C'est l'enjeu d'un programme de recherche BnF.

L'atelier-laboratoire de modélisation des collections numériques

Ce programme de recherche propose de concevoir des prototypes autour des collections numérisées de la BnF. Il s'agit d'expérimenter, avec une équipe de chercheurs et des étudiants de master en hypermédia, de nouveaux modes de médiations aux ressources numériques, dans le cadre d'un atelier-laboratoire IDEFI-CréaTic. La méthodologie proposée est celle de la recherche-création. L'objectif est de produire des objets éditoriaux innovants en développant des modes narratifs grâce à la visualisation de données et aux médias interactifs.

Ce programme poursuit comme objectif scientifique d'étudier et de proposer de nouvelles formes de médiations culturelles, scientifiques, territoriales et pédagogiques en utilisant les nouvelles technologies en Sciences de l'Information et de la Communication (SIC). Il s'agit de concevoir, créer et évaluer des dispositifs innovants dédiés à la transmission et à l'édition numérique. La problématique est celle de la « remédiatisation » des collections : au-delà de la numérisation des œuvres, quels modes d'accès proposer au public ? Quelles interfaces inventer face à l'accroissement exponentiel des ressources ? Comment naviguer dans des corpus complexes³ au-delà de la liste de résultats et du filtrage à facettes ?

Le programme s'insère dans le contexte des humanités numériques⁴ et ambitionne d'apporter des solutions du point de vue des interfaces, des usages et de la médiation. Il s'inscrit dans les perspectives de développement d'une offre de Gallica adaptée aux usages numériques du « grand public », conformément aux orientations stratégiques de la BnF.

3/ L'Odysée, matière exemplaire d'une remédiatisation

Pour expérimenter cela, l'*Odysée* apparaît comme une matière exemplaire. Première œuvre de la littérature occidentale (avec l'*Illiade*), c'est aussi une œuvre interculturelle qui, à travers le voyage d'Ulysse, fait le lien entre les rives de la Méditerranée. Géographie mythique des Anciens, elle s'ancre dans un contexte

³ Nous entendons par « corpus complexes » des ensembles pluridisciplinaires mêlant tout type de sources (textes, documents sonores, multimédia, images, films, photographies, données du web, métadonnées bibliographiques...).

⁴ Les humanités numériques (*Digital Humanities*) réunissent, autour de projets scientifiques innovants, différents acteurs de la recherche, au croisement de l'informatique et des arts, lettres, sciences humaines et sciences sociales.

historique dont elle s'affranchit par le principe même l'épopée, transfiguration poétique de l'espace et du temps. La matière de l'œuvre provient du fond des âges. Elle témoigne de « siècles obscurs » (Finley, 1954) autant d'une renaissance, celle de la civilisation grecque au VIII^e siècle avant notre ère, revivifiée par l'écriture et l'essor du commerce. De l'oral à l'écrit, l'œuvre, remédiatisée au cours des siècles, transcende ses différents supports pour s'offrir au lecteur, revisitée, régénérée dans de nouvelles formes. Aujourd'hui encore, elle imprègne notre imaginaire et touche un large public. L'*Odyssée*, monde fabuleux de dieux, de héros et de monstres, est aussi celui de lieux bien réels, d'objets archéologiques et d'œuvres artistiques.

Pour Gérard Genette, l'*Odyssée* est « une œuvre hypertextuelle, et, symbolique, la première en date que nous puissions pleinement recevoir et apprécier comme telle. » (Genette, 1982 : 246). Il entendait ainsi la profondeur intertextuelle d'une œuvre qui, multipliant récits et narrateurs, comprend ses propres réécritures et s'affirme comme un modèle d'écriture hypertextuelle, « qui fait un peu plus qu'autoriser les reprises ironiques, soupçonneuses, volontairement vertigineuses d'un Giraudoux, d'un Joyce, d'un Giono, d'un John Barth. » Chaque scène, chaque épisode donne lieu à des commentaires et des réécritures : pas moins de 96 références identifiées pour la seule escale d'Ulysse chez Circé (Escola et Rabau, 2015).

L'*Odyssée* offre ainsi l'exemple d'un livre-réseau, un « livre discuté, commenté, repris ; le livre générateur d'autres livres » : « De l'*Odyssée* d'Homère à l'*Ulysse* de Joyce, en passant par les poèmes dits « homériques », l'*Enéide* de Virgile et la *Franziade* de Ronsard ou les *Elégies* de Properce, nous avons sous les yeux un exemple canonique de la manière dont la littérature fonctionne : par reprises et gloses incessantes, commentaires et interprétations, adaptations à des contextes culturels variés qui témoignent de son infinie fécondité. » (Mounier, 2010).

À travers ses lectures allégoriques, l'*Odyssée* s'affirme comme un archétype de « l'œuvre ouverte » (Eco, 1962). Elle apparaît doublement ouverte puisque son texte ne pas connaît pas de forme achevée par un auteur : pas d'œuvre originale mais, au contraire, de multiples variations et variantes. Pour Umberto Eco, c'est un hypertexte de l'*Odyssée* qui reste « l'œuvre la plus ouverte dont il nous soit permis de parler » : *Ulysse* de James Joyce.

Textes, hypertextes et métatextes, l'œuvre se trouve ainsi au

centre d'un réseau⁵. Ne constitue-t-elle pas à elle seule une bibliothèque et un musée à travers les âges ?

Les représentations picturales relèvent aussi d'une forme de « lecture » du texte. Dès le VII^e siècle av. J.-C., de très nombreux vases grecs illustrent des scènes de l'*Odyssée*. Fresques et tableaux de l'Antiquité ont hélas disparu. Restent les vases et les sculptures, ainsi que les nombreuses copies romaines. Les peintres modernes ont également été inspirés par l'*Odyssée*, dès la Renaissance, s'attachant tout particulièrement aux figures féminines, dont ils amplifient la polysémie.

L'*Odyssée* nous apparaît ainsi comme la somme de ses lectures et relectures : qu'elles soient poétiques, linguistiques, artistiques, allégoriques, géographiques, historiques ou anthropologiques, celles-ci constituent pour nous comme des couches ou des plateaux : ce sont les épaisseurs du sens de l'œuvre.

Peut-on saisir ces multiples aspects dans une interface innovante qui remédiatiserait l'*Odyssée* à travers les collections patrimoniales de la BnF et témoignerait ainsi de la profondeur de l'œuvre ?

Il ne s'agit pas de concevoir une édition augmentée qui serait exhaustive, totale, mais de restituer la richesse de l'œuvre et sa complexité à travers une « expérience utilisateur » constituant une médiation au corpus odysseén, un pont entre lecture loisir et lecture savante. L'enjeu est d'inventer une forme nouvelle qui donne accès aux formes antérieures, tout en proposant de nouveaux modes de lectures et d'appropriation. Pour cela, nous adoptons une posture pragmatique articulant recherche et création.

La méthode : recherche-création en atelier-laboratoire

L'approche méthodologique est celle de la recherche-création dans le cadre d'un atelier-laboratoire IDEFI-CréaTic⁶. Cette

⁵ Une interface pédagogique a tenté de modéliser ce réseau sous la forme d'une constellation : <http://www.surgissantes.com/constellation?id=018>

⁶ Initiative d'excellence en formation innovante, IDEFI-CréaTic est un programme qui réunit des masters pluridisciplinaires sous la forme d'ateliers-laboratoires valorisant la pédagogie par projet.

démarche postule la création comme moteur épistémologique par la mise en œuvre d'une pédagogie de projet adossée à des fondements théoriques. L'atelier-laboratoire se présente comme un lieu de transversalité où sont associés chercheurs et professionnels autour des questions de médiation. Notre objectif est d'interroger les nouvelles formes et nouveaux usages du livre en abordant par l'expérimentation la problématique du « livre augmenté » dans la perspective de remédier et d'éditorialiser l'*Odyssee*. Trois axes de travail ont été privilégiés :

- **le corpus** : articuler, à partir d'un texte de référence, une sélection d'œuvres et de documents autour de la transmission du texte et de son interprétation (art, littérature, etc.) ;
- **le voyage** : représenter le voyage d'Ulysse sur une carte interactive et poser la question de la « géographie odysseenne » à partir des travaux de l'ethnologue Jean Cuisenier (2003) ;
- **le récit** : découvrir les multiples facettes de l'œuvre, sa dimension symbolique, initiatique, et ses réécritures.

L'atelier-laboratoire s'est déroulé au second semestre 2015-2016 en trois temps. Le premier temps fut consacré à la recherche et aux fondements théoriques avec un séminaire lors duquel une dizaine d'invités ont abordé l'*Odyssee* et l'édition augmentée sous différents aspects scientifiques, technologiques ou éditoriaux. Le deuxième temps fut celui du projet tutoré, encadré par l'équipe pédagogique du master Création et éditions numériques (CEN) de Paris 8 : les étudiants, mis en posture de concepteurs et d'utilisateurs, ont travaillé en mode projet afin de réaliser maquettes et prototype selon une méthode itérative passant progressivement des phases de conception et d'éditorialisation à celle du design et du développement. Le troisième temps fut celui de la transposition de l'*Odyssee*, en tant que « concept », dans le monde contemporain, et de l'appropriation de la figure d'Ulysse dans un exercice d'écriture hypermédiatique, réalisé en immersion à Athènes.

Le prototype : un « théâtre des errances d'Ulysse »

Le géographe grec Strabon parlait ainsi du voyage d'Ulysse à travers la Méditerranée. Bien entendu, il ne s'agit pas de dire que tel fut l'itinéraire du héros, qui reste une fiction, mais de s'appuyer sur la tradition de la « géographie odysseenne » comme hypothèse de travail et principe de médiation, afin de

permettre au lecteur de suivre les traces archéologiques et fictionnelles d'Ulysse. Trois accès sont proposés dans l'œuvre :

- une « **carte interactive** » retrace le voyage d'Ulysse d'après les hypothèses de Victor Bérard et permet d'accéder aux différentes étapes des aventures du héros ;
- une « **bande défilée** », développée en parallaxe, permet de suivre le récit d'Ulysse mis en scène par les œuvres patrimoniales de la BnF afin d'appréhender la profondeur historique et artistique de l'*Odyssée* à travers quelques unes de ses représentations au cours des âges ; la légende et le commentaire des œuvres s'affichent au clic ou au tap sur l'icône (+) ; des lectures d'extraits du texte ponctuent le parcours ;
- une « **bibliothèque** », celle des Sirènes ou de Circé, permet d'accéder à des ressources complémentaires, iconographiques et textuelles, organisées dans des parcours de lecture thématiques, autour des arts et traditions, interprétations et réécritures.

Le prototype fonctionnel a été conçu et réalisé par les étudiants du master Création et édition numériques (CEN) de Paris 8, dans le cadre de l'atelier-laboratoire IDEFI-CréaTic « Édition numérique augmentée », avec le soutien de la Bibliothèque nationale de France (BnF) et de la chaire Unesco ITEN (Innovation, Transmission et Édition Numérique). Ce travail s'inscrit dans le cadre du programme de recherche « Modélisations des collections numériques » (DED-LABO) de la BnF, qui vise à proposer, à travers des interfaces innovantes, des médiations « grand public » aux collections de la bibliothèque numérique Gallica. Il a bénéficié de la collaboration d'Aurélien Berra et des étudiants du master Humanités classiques et Humanités numériques (HCHN) de Paris 10. Le laboratoire NewTech (NTLab Uda), de l'Université d'Athènes, est également partenaire du projet.

En conclusion : modéliser le réseau

Le livre est un « nœud dans un réseau » : la formule de Michel Foucault nous a guidée tout au long de notre réflexion théorique et dans notre approche expérimentale du livre augmenté. Notre objectif est bien de proposer une interface au nœud et non de modéliser le réseau, car une telle ambition dépasserait largement le cadre de notre atelier-laboratoire. Cependant, nous pouvons avancer en conclusion la vision prospective Robert Darnton qui

ouvre des pistes de travail. Darnton imagine le livre électronique organisé en « six couches pyramidales », depuis la couche supérieure qui se contenterait d'un exposé concis du sujet jusqu'à la sixième couche qui serait le lieu d'un forum de discussion. Au sommet, « un exposé concis du sujet, peut-être déjà disponible en édition courante ». La couche suivante serait une série de versions développées de divers aspects « qui nourrissent le récit du registre supérieur ». La troisième couche « rassemblerait des documents » commentés. Une quatrième couche pourrait être théorique ou rétrospective avec des « extraits et des analyses des travaux antérieurs ». Une cinquième couche pourrait être pédagogique avec des supports de cours. Une sixième couche pourrait réunir des « rapports de lecture », les correspondances avec les lecteurs ; cette dernière couche s'élargirait avec le temps. (Darnton, 2011).

Cette approche articule les dimensions discursive, computationnelle et réticulaire que nous avons exposées, dans une lecture en plateau propre au rhizome dont la structure est une fractale qui, précisément, se présente comme un agencement de pyramides imbriquées. La pyramide est-elle l'ultime vision du livre augmenté ?

Références bibliographiques

Baccino, Thierry, 2004. *Mesure de l'utilisabilité des interfaces*. Paris, Lavoisier, Hermès sciences.

Berra, Aurélien, 2012. « Faire des humanités numériques », dans Mounier, Pierre (éd.), *Read/Write Book 2. Une introduction aux humanités numériques*, Marseille, OpenEdition Press, pp. 25-43. En ligne : <http://press.openedition.org/238>

Bolter, Jay David, et Grusin, Richard, 1999. *Remediation. Understanding New Media*. Cambridge, MIT Press.

Chartier, Roger, 2009. « La mort du livre ? », *Communication & langages*, pp. 57-65.

Chartier, Roger, 2012. « Qu'est-ce qu'un livre ? », *Écrit et cultures dans l'Europe moderne*, cours au Collège de France. En ligne : https://www.college-de-france.fr/media/roger-chartier/UPL62059_Chartier.pdf

Cuisenier, Jean, 2003. *Le périple d'Ulysse*. Paris, Fayard.

Darnton, Robert, 2011. *Apologie du livre : demain, aujourd'hui, hier*. Paris, Gallimard.

Dacos, Marin, et Mounier, Pierre, 2010. *L'édition électronique*. Paris, La Découverte.

Deleuze, Gilles, et Guattari, Félix, 1980. *Mille plateaux*. Paris, Éditions de Minuit.

Eco, Umberto, 1962. *L'œuvre ouverte*. Paris, Seuil.

Escola, Marc, et Rabau, Sophie, 2015. *Littérature seconde ou la bibliothèque de Circé*. Paris, Kimé.

Finley, Moses I., 1954. *Le monde d'Ulysse*. Rééd. Paris, Le Seuil, « Points-Histoire », 2002.

Foucault, Michel, 1969, *L'archéologie du savoir*, Paris, Gallimard.

Genette, Gérard, 1982. *Palimpsestes. La littérature au second degré*. Paris, Le Seuil, rééd. Points-Essais, p. 246-247.

Goody, Jack, 1979. *La raison graphique. La domestication de la pensée sauvage*. Paris, Éditions de Minuit.

Grafton, Anthony, 2012. *La page de l'Antiquité à l'ère du numérique. Histoire, usages, esthétiques*. Paris, Hazan / Musée du Louvre.

Laborderie Arnaud, « Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica », in *Documents et dispositifs d'information à l'ère post-numérique*, actes du 18^e colloque international sur le document électronique, CiDE.18, Paris, Europa, 2015.

En ligne : <https://hal-bnf.archives-ouvertes.fr/hal-01239425>

Moretti, Franco, 2008. *Graphes, cartes et arbres. Modèles abstraits pour une autre histoire de la littérature*. Paris, Les Prairies ordinaires.

Mounier, Pierre, 2010. « Le livre et les trois dimensions du cyberspace », in Dacos, Marin (dir.), *Read/Write Book. Le livre inscriptible*, Marseille, OpenEdition Press (p. 175-183). En ligne : <http://books.openedition.org/oep/179>

Pignier, Nicole, et Drouillat, Benoît, 2004. *Penser le web design. Modèles sémiotiques pour les projets multimédias*. Paris, L'Harmattan.

Sinatra, Michael E., et Vitali-Rosati, Marcello, 2014. *Pratiques de l'édition numérique*. Montréal, Presses de l'Université.

Ecosystèmes de connaissances pour les archives numériques

Samuel Szoniecky

Laboratoire Paragraphe – CiTU

Dans des domaines aussi différents que l'innovation en entreprise, les archives personnelles, l'archéologie ou les archives audiovisuelles, l'intelligence collective est à l'œuvre pour donner du sens au chaos numérique qui ne cesse de croître avec la prolifération des traces, de leur conservation et des recherches en humanités numériques. Quels sont les rapports qu'entretiennent ces intelligences collectives avec les archives numériques ? Comment utiliser au mieux les technologies intellectuelles pour aider les humains dans leurs compréhensions du numérique ? Existe-t-il des outils et des méthodes qui mettent à profit la raison computationnelle [PAR 16] sans pour autant négliger l'éthique humaine ?

Pour répondre à ces questions nous analyserons en quoi ces intelligences collectives mises au service des archives numériques constituent des écosystèmes de connaissances. Nous montrerons comment analyser les existences qui peuplent ces écosystèmes suivant une méthode générique de modélisation qui associe description ontologique et potentialité éthique afin de proposer aux chercheurs en humanités numériques des outils pratiques et des méthodes opérationnelles.

Modélisation génériques des écosystèmes de connaissances

Les dispositifs informationnels que nous utilisons ne consistent plus à figer dans le marbre ou le papier nos connaissances. En passant par le numérique, ces dispositifs ont définis de nouvelles frontières [SAL 14] et de nouvelles formes d'information qui

s'apparentent davantage à celles des êtres vivants plutôt que celles des objets inertes. A la fois dynamiques et interactives, ces informations évoluent sans cesse en interactions avec d'autres informations suivant des processus réalisés par des acteurs algorithmes et/ou humains.

Pour décrire et analyser ces dispositifs, les informations et les connaissances qui leurs sont liées, nous privilégions l'analogie [HOF 13] de l'écosystème à celle de l'architecture qui véhicule des présupposés liés à une activité éditoriale [LYP 13]. Une méthode de modélisation qui à l'ambition d'être générique doit pouvoir traiter des informations qui ne respectent pas forcément les impératifs éditoriaux de cohérence, de cohésion et de sens commun. Des dispositifs informationnels peuvent être incohérents, mal-conçus voir même complètement hermétique à tout compréhension par exemple dans le cas d'œuvres poétiques. La modélisation écosystémique ouvre l'univers des possibles en incluant tout type de connaissances, d'informations et de dispositifs. Surtout, elle prend en compte les dimensions vivantes, évolutives et interactives des éléments qu'elle modélise en incluant par exemple des états de croissance, de maturité, de mort comme des URL qui ne sont plus valides.

Cette méthode s'articule autour de quatre objectifs principaux. Le premier objectif est de modéliser des arguments sous la forme d'agents autonomes [FER 97] possédant leur propre mode d'existence [LAT 12] et pouvoir d'agir [RAB 05]. Le deuxième objectif consiste à générer automatiquement l'évolution de ces arguments-existences pour produire des controverses. Le troisième objectif consiste à rendre interopérable les analyses de controverses. Le quatrième objectif à l'ambition de faire émerger un écosystème de connaissance à partir des interactions entre arguments, chercheurs et controverses.

Description ontologique des arguments

De notre point de vue, un argument est une existence dans un écosystème de connaissance. La description ontologique des arguments passe selon cette hypothèse par la modélisation d'existences informationnelles composées de quatre types d'éléments : des documents, des acteurs, des concepts et des rapports entre ces éléments. Nous travaillons à la conception d'une application Web qui a pour objectif de guider les utilisateurs dans ce processus de modélisation en leur donnant

une représentation graphique de cette description. Nous ne détaillerons pas ici ce travail qui a déjà été présenté [SZO 15].

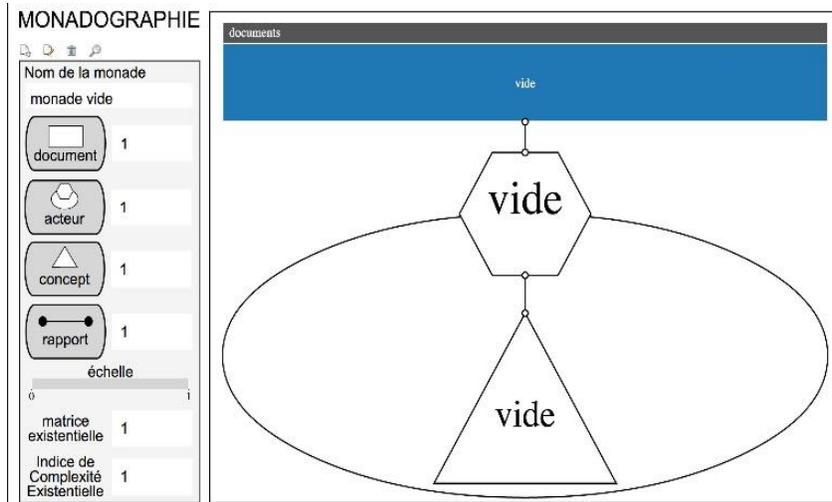


Figure 1 : modélisation générique d'une existence informationnelle

Simulation des controverses

L'objectif de cette simulation est de fournir des hypothèses heuristiques pour explorer et modéliser un écosystème de connaissances. La simulation des controverses consiste à générer à partir des existences-arguments de nouveaux éléments pour évaluer leur viabilité et dans ce cas transformer les controverses en arguments pour le cas échéant s'en servir pour simuler de nouvelles controverses. Ainsi, ce met en place un cycle d'interprétation dont on pourra suivre chronologiquement les étapes d'exploration (controverses) et de modélisation (arguments).

Pour générer ces cycles interprétatifs nous exploiterons le potentiel génératif des existences suivant deux dynamiques liées à une augmentation ou à une diminution du nombre d'éléments de chaque type.

Dynamique d'augmentation

La dynamique d'augmentation s'applique soit sur un élément unique soit sur plusieurs éléments de différents types.

Dans le cas d'un seul élément, la croissance se déroule suivant deux mouvements complémentaires : interne et externe. Ces mouvements peuvent être qualifiés suivant les principes de sémantique différentielle [BAC 07 p. 142] qui définissent les relations de communauté ou de différence suivant la position relative d'un élément par rapport à deux axes, celui du père et du frère dans un arbre. Dans ce cas, le mouvement de croissance interne consiste à générer des relations de différences avec les frères à partir d'une relation de communauté avec le père. Par exemple un document « livre » (communauté-père) pourra croître intérieurement en générant des documents « chapitres » (différences-frères) eux-mêmes pouvant être les pères se décomposant en phrases, etc. Autre exemple, un acteur « laboratoire de recherche » (communauté-père) aura une croissance interne en générant les acteurs « chercheurs » (différences-frères). Pour un seul élément, la croissance externe correspond à la génération d'une relation de communauté avec le père, par exemple dans le cas de l'acteur « laboratoire de recherche », une croissance externe peut être générée en créant une relation communauté-père avec l'acteur « université ».

Dans le cas de plusieurs éléments de différents types ou seul le mouvement de croissance externe s'applique, la sémantique référentielle n'est pas suffisante pour modéliser les relations entre les éléments car celles-ci ne se déploient pas uniquement suivant le modèle de l'arbre mais aussi suivant le modèle du rhizome [DEL 80, p. 31]. Même si on peut utiliser les relations de type différence-père et communauté-frère pour caractériser les relations entre des éléments du même type, les possibilités de relations sont beaucoup plus nombreuses. Le modèle générique de stockage des données historiques du projet SyMoGIH [BER 12] propose une liste non exhaustive des « unités de connaissance » qui donne un aperçu des relations possibles entre deux éléments.

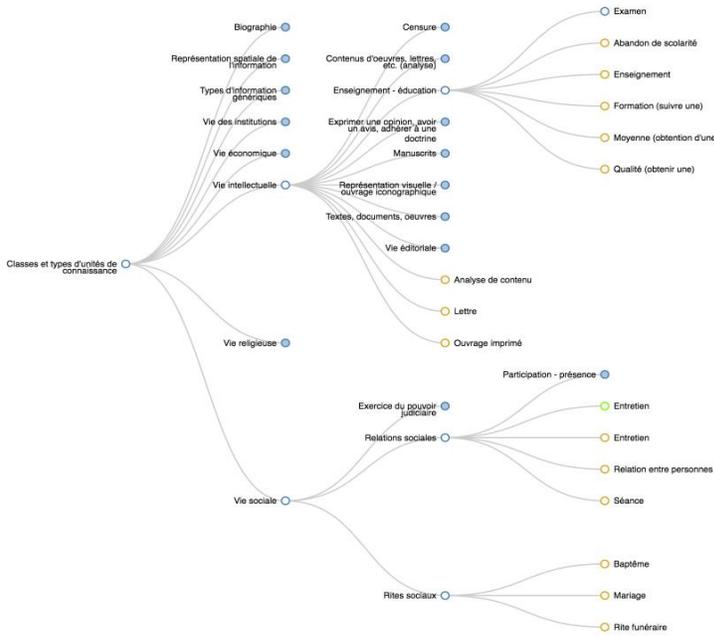


Figure 2 : arborescence des unités de connaissance SyMoGIH

A partir de la complémentarité des mouvements de croissance élémentaires d'un élément (interne-externe) et des plusieurs (externes), on pourra définir une collection d'algorithmes génériques pour faire croître chaque type d'élément ou d'existence. Nous avons déjà donné deux exemples avec le document-livre qui se décompose en document-chapitre ou l'acteur-chercheur qui est mis en relation avec les autres acteurs-chercheurs d'un même acteur-institution. Dans l'un ou l'autre des cas, les algorithmes peuvent être employés de manière récursive pour appliquer la génération des éléments et des relations à chaque élément nouvellement créé. Par exemple, nous pouvons appliquer à un livre l'algorithme « concept-matière Rameau (communauté-père) => document-livre (différences-frères) » pour générer les documents-livres ayant les mêmes matières Rameau puis pour chaque document-livre créé, appliquer ce même algorithme pour générer le réseau des documents-livres et des concepts-matières Rameau.

Il est très important de noter que les algorithmes génératifs ne sont pas nécessairement complètement automatique et qu'ils peuvent avoir besoin d'action humaine, par exemple pour créer une relation entre un acteur-lecteur, un document-phrased et un concept-émotion.

Dynamique de diminution

L'usage des algorithmes génératifs entraîne la prolifération des éléments ce qui rend difficile voire impossible leurs analyses. Parallèlement à la dynamique de croissance, la dynamique de diminution permet d'éviter la surcharge cognitive en rendant plus simple les existences informationnelles. Surtout, cette dynamique permet de simuler de nouvelles hypothèses heuristiques en proposant des niveaux d'échelle différents ou de nouvelles controverses par simplification des arguments. La dynamique de diminution des éléments suit elle aussi deux mouvements celui de concentration et celui de suppression.

Le mouvement de concentration consiste à regrouper des éléments de même type. Ce mouvement reprend lui aussi les principes de la sémantique différentielle. Par exemple, plutôt que d'afficher tous les documents-phrases (communauté-frère) d'un document-livre (communauté-père), on diminue ces éléments en affichant uniquement le document-livre. Chaque type d'élément étant autonome les uns par rapport aux autres, on pourra s'interroger sur la pertinence des rapports entre le document-livre et les concepts avant et après une concentration : est-ce que les concepts utilisés pour décrire les document-phrases restent pertinent au niveau du document-livre ? Ce type d'interrogation peut être très utile pour paramétrer les outils de traitement automatique de la langue qui génèrent souvent des relations conceptuelles qui peuvent être pertinentes au niveau d'une phrase mais que ne le sont plus au niveau d'un livre et encore moins au niveau d'un corpus.

Le mouvement de suppression consiste à éliminer un élément d'une existence-argument afin de la simplifier et de proposer une nouvelle forme de controverse. Il faut noter que la suppression d'un élément entraînera la suppression des éléments avec lesquels il est en rapport et donc de jouer sur la viabilité d'un argument. Ce mouvement de suppression est fondamental pour l'ingénierie des connaissances et le développement de technologies intellectuelles performantes car elle permet de modéliser des profils socio-sémantiques complexes qui prennent en compte ce qui n'intéresse pas un individu et donc de filtrer les informations non-pertinentes.

Interopérabilité des analyses

L'interopérabilité des analyses est sans doute l'objectif principal de la modélisation générique des écosystèmes de connaissances. Notre ambition est de fournir les technologies intellectuelles

pour organiser le « crowdsourcing » [AND 2016] des travaux d'exploration, de modélisation et d'interprétation qu'ils soient faits par des chercheurs, des étudiants ou de simples amateurs. Pour ce faire, il est nécessaire de transformer l'activité intellectuelle de ces personnes dans un langage symbolique qui soit calculable par des machines. Nous avons montré les difficultés d'une telle entreprise et comment les dépasser en rendant transparent la manipulation de ces langages symboliques par l'usage d'analogies graphiques [SZO 12].

Nous disposons aujourd'hui de plateforme de Linked Open Data (LOD) comme celles de la BNF⁷, de DBPEDIA⁸ ou de SyMoGIH dont les performances et la pérennité offre des références symboliques parfaitement utilisables pour ce type de projet. Parallèlement, le W3C travail à des spécifications pour l'annotation ouverte des documents⁹ qui fournit un cadre formel pour la description des rapports entre des documents, des acteurs et des concepts. En bref, les conditions techniques au niveau des langages symboliques sont aujourd'hui réunies pour développer des analyses interopérables.

⁷ <http://data.bnf.fr/>

⁸ <http://fr.dbpedia.org/>

⁹ <https://www.w3.org/annotation/>

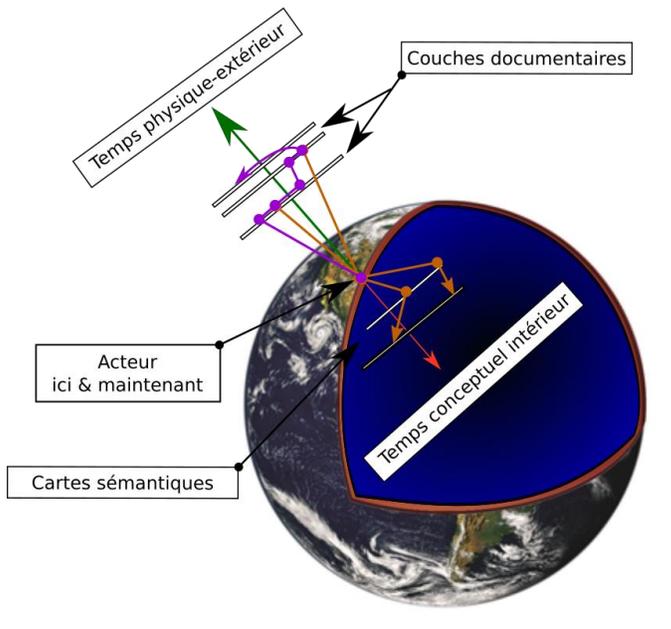


Figure 3 : pulsations existentielles

L'analogie de l'écosystème et plus particulièrement celle de l'arbre que nous utilisons pour modéliser les existences informationnelles offre une formalisation graphique évidente pour représenter les dynamiques d'augmentation et de diminution ainsi que les caractéristiques spatio-temporelles sous la forme de pulsations existentielles (Figure 3 : pulsations existentielles³³) [SZO 16]. De plus, le modèle de diagramme que nous proposons permet de calculer le potentiel génératif correspondant à la matrice des documents, des acteurs, des concepts et des rapports. Cette forme générique est aussi utile pour comparer la complexité des arguments et des controverses suivant leur indice de complexité [SZO 14].

Vers des écosystèmes de connaissances

Sans aller jusqu'aux visions prophétiques d'une intelligence collective trans-humanistes¹⁰ ni jusqu'à la science-fiction où les technologies intellectuelles font émerger des sociétés utopiques [HES 43] ou dystopiques [ORW 49], on peut envisager un futur

¹⁰ <http://humanityplus.org/philosophy/transhumanist-faq/>

du travail intellectuel bien plus efficace que ce qu'il est aujourd'hui. La méthode de modélisation des écosystèmes que nous présentons va dans ce sens en proposant les moyens de rendre interopérable l'exploration, la modélisation et l'interprétation des connaissances.

Il faut comprendre cette méthode logico-graphique comme une interface visuelle et conviviale pour la lecture et l'écriture de connaissances formalisées pour les traitements automatiques. De la même façon que les algorithmes de spatialisation des réseaux utilisés par GEPHI¹¹, les algorithmes génératifs dont nous avons parlé doivent être envisagés à travers une interface permettant de les concevoir et les paramétrer. Même si un gros travail reste à faire pour développer ces algorithmes et les consolider par l'expérience des utilisateurs, les pistes qui ont déjà été explorées [CHA 15] [LEV 16] nous encouragent à continuer ces efforts notamment dans le domaine de la modélisation des dynamiques scientifiques par des systèmes d'agents [CHA 16].

Conclusion

Cette proposition de méthode nous amène à concevoir des existences-arguments qui sont autant de « graines sémantiques », c'est à dire la formalisation générique d'une potentialité sémantique. Ces graines portent en elles une dynamique algorithmique pouvant générer des controverses qui guident l'usager dans le maelstrom cognitif des archives numériques en faisant pousser des controverses sous la forme de rapports entre des documents, des acteurs et des concepts. C'est à partir de cette pousse continue de potentialité sémantique que l'usager peut construire son interprétation en sélectionnant les rapports qui lui semblent adéquates pour ensuite partager le fruit de ces réflexions dans un langage formel calculable.

Existe-t-il un ADN des graines sémantiques ? Qui sont les créateurs des graines sémantiques ? Des ingénieurs sémanticiens ? Des cognitiens biologistes ? Peut-on valoriser les archives numériques comme on cultive un jardin ? A travers ces questions, ce qui est globalement en jeu c'est la capacité des interfaces homme-machine à rendre accessible la complexité des

¹¹ <https://github.com/gephi/gephi>

formalismes à des non-informaticiens sans perdre la complexité des expertises dans un domaine.

Bibliographie

[AND 2016] Andro, M. et Saleh, I., « Le crowdsourcing appliqué aux bibliothèques numériques », *Bulletin d'Informations- Association des Bibliothécaires Français*, n° 83, p. 23-25, 2016.

[BAC 07] Bachimont, B., *Ingénierie des connaissances et des contenus : Le numérique entre ontologies et documents*. Paris: Hermes science publications, 2007.

[BER 12] Beretta, F. et Vernus, P., « Le projet SyMoGIH et la modélisation de l'information : une opération scientifique au service de l'histoire », *Les Carnets du LARHRA*, n° 1, p. 81-107, 2012.

[CHA 16] Chavalarias, D., « What's wrong with Science? », *Scientometrics*, p. 1-23, août 2016.

[CHA 15] Chateauraynaud, F., « Trajectoires argumentatives et constellations discursives », *Réseaux*, n° 188, p. 121-158, févr. 2015.

[CIT 10] Citton, Y. *L'avenir des humanités : Economie de la connaissance ou cultures de l'interprétation*, Editions La Découverte, 2010.

[CIT 08] Citton, Y. "Entre l'économie psychique de Spinoza et l'inter-psychologie économique de Tarde", in: *Spinoza et Les Sciences Sociales : De La Puissance de La Multitude À L'économie Des Affects*. Éd. Amsterdam, Paris, 2008.

[DEL 68] Deleuze, G. *Spinoza et le problème de l'expression*. Paris: Éditions de Minuit, 1968.

[DEL 80] Deleuze, G. et Guattari, F., *Mille plateaux*. Paris: Éditions de minuit, 1980.

[DES 05] Descola, P., *Par-delà nature et culture*. Paris: NRF : Gallimard, 2005.

[FER 97] Ferber, J., *Les systèmes multi-agents : vers une intelligence collective, Informatique, intelligence artificielle*. InterÉd., Paris, 1997.

- [HES 43] Hesse, H., *Le Jeu des perles de verre*. Le Livre de Poche, 2002, première parution 1943.
- [HOF 13] Hofstadter, D. et Sander, E., *L'analogie : Coeur de la pensée*. Odile Jacob, 2013.
- [LAT 12] Latour, B., *Enquêtes sur les modes d'existence : Une anthropologie des Modernes*. Editions La Découverte, 2012.
- [LEV 11] Lévy, P., *La sphère sémantique : Tome 1, Computation, cognition, économie de l'information*. Hermes Science Publications, 2011.
- [LEV 16] P. Lévy, « L'intelligence algorithmique Exploration de la sphère sémantique », 2016, https://www.dropbox.com/s/pnoha3fqwp2dcgc/DEF-L%27INTELLIGENCE_ALGORITMIQUE.pdf?dl=0.
- [LYP 13] Lipsyc, C. et M. Ihadjadene, « Architecture de l'information et éditorialisation », *Études de communication. langages, information, médiations*, n° 41, p. 103-118, déc. 2013.
- [ORW 49] G. Orwell et A. Audiberti, *1984*, FOLIO. Paris: Gallimard, 1972, première parution 1949.
- [PAR 16] Parisi, L., « La raison instrumentale, le capitalisme algorithmique et l'incomputable », *Multitudes*, n° 62, p. 98-109, avr. 2016.
- [PRI 11] Prié, Y. *Vers une phénoménologie des inscriptions numériques. Dynamique de l'activité et des structures informationnelles dans les systèmes d'interprétation*. (HDR). Université Claude Bernard - Lyon I, Lyon, 2011.
- [RAB 05] Rabardel, P., "Instrument subjectif et développement du pouvoir d'agir", in: *Modèles Du Sujet Pour La Conception : Dialectiques, Activités, Développement*. Octarès éditions, Toulouse, 2005.
- [RAS 01] Rastier, F., Cavazza, M, Abeille, A. (dir.), *Sémantique pour l'analyse. De la linguistique à l'informatique*, Masson, Paris, 2001.
- [SER 09] Serres, M., *Ecrivains, savants et philosophes font le tour du monde*. Paris: Pommier, 2009.
- [SAL 14] Saleh, I., Bouhai, N., et H. Hachour, *Les frontières du numérique*. Paris: Editions L'Harmattan, 2014.
- [SZO 12] Szoniecky, S., « Évaluation et conception d'un langage symbolique pour l'intelligence collective : Vers un

langage allégorique pour le Web », Thèse de Science de l'information et de la communication, Université Paris VIII Vincennes-Saint Denis, 2012.

[SZO 14] Szoniecky, S. et Hachour, H., « Monades pour une éthique des écosystèmes d'information numériques », présenté à *Digital Intelligence*, Nantes, 2014.

[SZO 15] Szoniecky, S. et Louâpre, M., « Outillages numériques pour les humanités : cartographier des réseaux d'influences », in *ISKO - Magreb 2015*, Hammamet, Tunisie, 2015.

[SZO 16] Szoniecky, S., Saleh, I. et Laborderie, A., « Digital monads for a spatio-temporal modeling of urban humanities », in *City Temporalities*, Europa., Paris, 2016.

Le rôle de records management dans la gestion stratégique de l'information des organisations

Anna Lezon Rivière

Université Paris 8 Vincennes-Saint-Denis

Laboratoire Paragraphe

anna.lezon-riviere@univ-paris8.fr

Mots clés : records management, archivage, pratiques professionnelles, gestion stratégique de l'information, organisation, retour d'expérience, documents numériques

Introduction

En réduisant voire en éliminant les coûts de production de l'information les organisations ont facilité un accroissement d'information de faible qualité. La relative facilité de publication offerte par les dispositifs informationnels (traitement de texte, CMS, etc.) et la généralisation des messageries conduit, selon certains auteurs, aux stress et à l'anxiété (Heylighen, 2002). Les acteurs sont désormais confrontés à plus d'informations qu'ils ne peuvent réellement traiter (Heylighen, 2002), (Melinat, Kreuzkam, & Stamer, 2014). Dans ce contexte, la gestion stratégique de l'information (GSI), en premier lieu, devrait apporter la réponse à la question fondamentale : que voulons-nous faire avec l'information dans l'organisation ? (Davenport & Prusak, 1997, p. 28).

Le records management (RM) (AFNOR, 2008) est l'une des fonctions info-documentaire de l'organisation qui permet la bonne exécution des activités de celle-ci. Il contribue à la protection des droits de l'organisation et de ses membres. Le

RM prend une part importante dans la capitalisation de la mémoire-objet de l'organisation. En gérant ce capital informationnel, le RM constitue un sous-ensemble opérationnel et opératoire de la GSI.

Dans cet article nous tentons de cerner le rôle du RM dans une gestion globale et intégrée de l'information au sein de l'organisation que nous nommons la gestion stratégique de l'information. Pour ce faire, cette étude exploratoire s'appuie sur la littérature ainsi que sur des entretiens que nous avons menés auprès des acteurs d'institutions publiques et de grandes entreprises françaises au sujet des politiques et des enjeux de RM et de l'archivage au sein des organisations.

GSI : complexité de la conceptualisation

Depuis près de cinquante ans, la gestion de l'information apparaît comme un domaine de recherche fécond en sciences de gestion et en sciences de l'information (Maceviciute & Wilson, 2002). Ce champ de recherche reste encore jeune et manque d'homogénéité dans les approches des concepts et des objets de recherche (Nowé Hedvall, 2007). Macevičiūtė et Wilson (2002) démontrent que le consensus sur la compréhension commune est loin d'être atteint aussi bien du point de vue disciplinaire qu'organisationnel. Aussi, le paradigme de la « gestion de l'information » apparaît encore flou (Wilson, 2002), (Madsen, 2013) et peut revêtir différents syntagmes et définitions. Nous pouvons citer, pour en donner l'exemple, les définitions de la norme AFNOR (AFNOR, 2004) de management et de la gestion de l'information. Le management de l'information correspond selon ce texte aux « *activités coordonnées pour orienter et contrôler un organisme en matière de gestion de l'information* » et la gestion de l'information est un « *processus couvrant le cycle de vie de l'information* ». En outre, nous pouvons noter la rigueur de nos collègues canadiens dans l'emploi des termes : « gestion de l'information » en langue française et « *information management* » dans les écrits en langue anglaise.

Le concept de la gestion « stratégique » de l'information est issu du mouvement de la gestion des ressources informationnelles et s'enrichit avec les courants de la gestion des connaissances et de la veille stratégique (Bergeron, Dufour, Maurel, & Mercier, 2010). Bergeron et ses pairs définissent la GSI en termes suivants : « *A la base se trouve la reconnaissance de*

l'information comme ressource qui doit être gérée, au même titre que les ressources humaines, matérielles ou financières, dans les organisations. C'est fournir la bonne information à la bonne personne au bon moment par une approche intégrée et stratégique ».

Nous pouvons enrichir cette réflexion à la lumière de nos travaux en gestion stratégique de l'information et en construction de sens (Lezon Rivière, 2013). Les résultats de nos travaux démontrent que la complexité de la problématique ne réside pas nécessairement dans « *la reconnaissance* » du rôle de l'information comme moyen vital de l'organisation mais dans sa mise en œuvre et dans les méthodes qui l'accompagnent (Lezon Rivière, 2013). « *L'écrasement* » du temps, l'instantanéité dans la coordination des actions qui se déroulent à partir des espaces géographiques et informationnels multiples, les technologies avancées contribuent à la complexification de la démarche de gestion stratégique de l'information au niveau organisationnel.

De même, la formulation « *fournir la bonne information à la bonne personne au bon moment* » ne nous explicite pas l'essentiel de la gestion stratégique de l'information : qu'est-ce « *la bonne* » information ? En puisant dans les approches *sensemaking* (Dervin, Weick), nous pourrions avancer que « *la bonne* » information est « *le sens en construction* » dans un espace et un temps définis permettant de déclencher une action jugée non-équivoque par l'acteur.

Dans notre approche, la GSI est dotée d'une vision qui se manifeste par ses temporalités, ses espaces informationnels, son historicité ainsi que la construction de sens individuel et collectif. Au sein de l'organisation, elle se dote d'une stratégie informationnelle holiste. Elle comprend la gestion des données, de l'information et des connaissances. Elle inclue la gestion de l'information de nature variée, notamment, interne et externe, « ordinaire » (routinière par exemple) et « stratégique ». En conséquent, cette approche considère toutes les ressources informationnelles de l'organisation ainsi que les technologies qui les supportent. La GSI fédère en son sein les démarches, notamment, de la gestion des connaissances, de la veille, de l'intelligence économique et du records management. Elle se préoccupe des interdépendances (apports mutuels, liens, redondances, etc.) entre ces fonctions pour en dégager une vision d'ensemble en lien avec d'autres grandes fonctions et métiers de l'organisation. Cette vision de la GSI s'inscrit dans la

stratégie de l'organisation. De ce fait, les acteurs majeurs de la GSI participent aux instances de la gouvernance de l'organisation.

RM : origines, pratiques, enjeux

Au début des années 1950, l'administration américaine est confrontée à une masse importante de documents accumulés durant le conflit (AAF, 2007, p. 62). Le président Truman met en place une commission dirigée par le sénateur Hoover pour résoudre la problématique de la gestion de cette masse documentaire. La commission émet une règle fondamentale qui est celle de la distinction entre les dossiers utiles aux affaires (*records*) de l'organisation et les archives, dont l'organisation n'a plus d'utilité mais, qui intéressent la collectivité dans leur aspect du témoignage historique. Peu après, en 1956, l'archiviste Théodore Shellenberg formalise la théorie « du cycle de vie » dans son ouvrage « *Modern Archives: Principles and Techniques* ». En France, Yves Pérotin (1961) formalise le concept de « trois âges » et propose les termes des « archives courantes », « archives intermédiaires » et « archives archivées » (historiques). Cette théorie, largement répandue, reste une référence de la pratique archivistique et du RM. Bien plus tard, en 1996, l'AFNOR a été saisie de la problématique de « gestion des enregistrements » suite au dépôt de la norme australienne AS 4390 auprès de l'ISO pour en faire un standard international (Drouhe, Keslassy, & Morineau, 2000, p. 3).

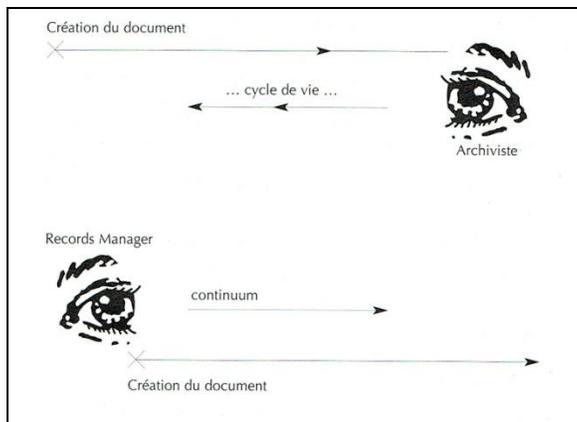


Fig. n°1 Changement d'axe du regard sur le document créé (Drouhe, Keslassy, & Morineau, 2000, p. 19)

La principale différence de records management avec l'archivistique est le changement du regard porté sur le document créé et son cycle de vie (Fig.n°1). Les pratiques archivistiques des siècles précédents tout autant que les formations dispensées avaient pour effet qu'« *aucun de nos anciens ne fut (...) sensible à la gestion des documents vivants* » (Dhérent, Sibille de Grimouard, Ott, & Brouland, 2009, p. 33). L'évolution de ces pratiques s'amorce considérablement dans les années 1970 où les archivistes interviennent activement sur les archives intermédiaires. L'avènement de l'ère numérique bouleverse les concepts établis et, en conséquence, les pratiques archivistiques. Les notions des « archives intermédiaires » et des « archives courantes » appliquées à la gestion des documents numériques perdent de leur pertinence. Dès lors, l'intervention *a posteriori* dans le cycle de vie « *n'est plus envisageable pour les documents numériques* » (Dhérent, Sibille de Grimouard, Ott, & Brouland, 2009, p. 34). Les pratiques de la gestion des documents dès la création des documents d'archives sont ainsi progressivement reconnues et intégrées dans les pratiques archivistiques avec une mention de spécificité accordée par les Archives de France en 2007 au Records Management. En France généralement, les pratiques de records management prennent d'abord ancrage dans le secteur privé avant de gagner le secteur public soumis de plus en plus aux phénomènes telles la judiciarisation de la société, la pression temporelle et la surabondance de l'information.

En effet, le records management « capture » le document à sa source et envisage « le continuum » de sa vie opérationnelle à partir de sa création. Le RM peut être considéré comme méthode de la gestion de la mémoire documentaire de l'organisation qui a pour objet « *l'ensemble des documents « à archiver », c'est-à-dire les documents que l'organisme aura décidé de préserver à titre de preuve ou en raison de leur valeur informationnelle* » (AAF & ADBS, 2005). De ce fait, le périmètre de RM, dans de nombreux cas, ne concerne pas tous les documents de l'organisation mais la sélection opérée selon les principes, notamment, de la preuve juridique, judiciaire ou règlementaire. Au sein de ce fonds, le RM désigne les documents vitaux de l'organisation que nous pourrions qualifier de stratégiques. Il s'agit des documents dont la perte ou la destruction entrainerait des risques élevés pour l'organisation voire sa disparition. Aussi, le RM « *a pour finalité de permettre à l'organisme de disposer à tout instant du document dont il a besoin pour*

conduire ses activités, répondre aux exigences légales et réglementaires, et se protéger en cas de contentieux » (AAF & ADBS, 2005, p. 4). Le système records management permet de tracer et d'identifier les responsabilités dans la gestion de ces documents. Il prend une part importante dans la capitalisation de la mémoire organisationnelle. Il la préserve, se préoccupe de sa sécurité physique aussi bien qu'informationnelle. Au-delà de cette fonction patrimoniale, le passage vers le numérique positionne le RM au carrefour de la gestion des risques, de la confiance, de l'efficacité personnelle et organisationnelle dans le travail et, par conséquent, de la productivité de l'organisation. Aussi, « *savoir créer des documents répondant à des exigences standardisées et les mettre en place dans des systèmes, savoir mettre en œuvre les processus de décision, évaluer et surveiller cette production* » (Cottin, 2015) apparaissent comme de nouvelles pratiques professionnelles à assimiler.

En effet, le format électronique des documents d'archives est de plus en plus natif et unique. Les organisations du secteur tant privé que public dématérialisent massivement les processus de travail et les documents qui les accompagnent. Les lois nationales et européennes viennent confirmer, encourager voire imposer cette propension. Dans ce contexte, la question de la mise en œuvre et de la gestion des systèmes électroniques de records management devient stratégique. L'étude des ERMS¹² par Gunnlaugsdottir (2006, 2008) auprès de huit organisations islandaises nous indique les objectifs de la mise en œuvre d'un tel système, formulés par les managers : augmenter la productivité, améliorer le service client, anticiper l'économie des coûts, réduire les espaces de stockage des documents, ingérer les procédures de travail, avoir un meilleur aperçu global des dossiers et enfin, mieux gérer les obligations légales (Gunnlaugsdottir, 2006, p. 235). Si les ERMS sont correctement implémentés et s'ils sont largement utilisés alors ils deviennent des outils puissants de l'échange, de la circulation de l'information, du travail collaboratif et de la gestion des connaissances au sein de l'organisation (Gunnlaugsdottir, 2008, p. 303).

Cependant, dans les organisations d'aujourd'hui, les documents d'archives se présentant dans différents formats (photos, vidéo, Web) sont véhiculés par de nombreux médias et se retrouvent

¹² *Electronic Records Management Systems*

sur différentes plates-formes techniques. C'est ainsi que l'étude de Williams et Ashley (2009) démontre que 78% des records managers interviewés n'avaient pas instauré des pratiques de conservation des sources en provenance du Web 2.0 tels les blogs, les pages Web ou des messages instantanés (Joseph, Debowski, & Goldschmidt, 2012, p. 68). Si nous considérons ce champ élargi de la gestion de l'information qui englobe de multiples business-processus, applications ou dépôts de données, il devient difficile voire hasardeux de se limiter à la seule démarche d'ERMS. Désormais, il apparaît vital, pour les organisations et les records managers, de se préoccuper de la gestion de l'ensemble des ressources informationnelles d'une organisation dont elle a besoin (Joseph, Debowski, & Goldschmidt, 2012).

Ce constat positionne le RM comme partie intégrante de la gestion stratégique de l'information au sein des organisations. Cela implique également le changement du rôle de records managers qui serait désormais amenés à réfléchir, avec d'autres acteurs, sur cette globalité de la démarche de la gestion de l'information. Trois enjeux principaux sont identifiés pour les professionnels de records management : 1) élargir le champs d'action à l'ensemble des ressources informationnelles de l'organisation, 2) mieux comprendre les attentes et les habiletés des utilisateurs pour proposer des outils simples, orientés usager sans se centrer sur le système, 3) être proactif dans l'exploitation des technologies pour automatiser les processus RM et, ainsi, soutenir les pratiques « de la connaissance » (Joseph, Debowski, & Goldschmidt, 2012, p. 69).

RM au sein des organisations : applications et retours d'expérience

Notre expérience professionnelle nous a amené à rencontrer de nombreux acteurs du RM et de l'archivage au sein de grands groupes et des institutions, principalement dans le contexte des études comparatives. Ces dernières visaient également le recueil du retour d'expérience, élément précieux durant la phase de la conceptualisation (AFNOR, 2008) du système RM. Nous avons également interviewé deux d'entre eux dans notre rôle de chercheur. Nous présentons ci-dessous trois cas d'études : un grand groupe de télécom, une institution culturelle et une institution de l'administration centrale.

Le groupe Orange est le premier exemple de l'application de RM que nous exposons. Cette entreprise est l'un des principaux opérateurs de télécommunication dans le monde. Elle est présente sur les cinq continents et compte 156 000 collaborateurs. Nous y avons rencontré deux records managers successifs : le premier en 2008 et le second en 2016. Le premier records manager conduisait le projet de records management au sein de l'entreprise pour améliorer la gestion de la documentation interne papier et électronique du groupe dans l'objectif d'une meilleure maîtrise des flux informationnels. L'analyse des risques était un des axes forts dans ce projet. Il s'agissait d'identifier la documentation stratégique et vitale pour l'entreprise pour en déceler les risques, possibles et probables, provoqués par la perte de cette documentation. L'ensemble des documents ont été pris en compte dans ce projet (les documents à valeur juridique aussi bien qu'informationnelle). Cette démarche a facilité la recherche d'information et a permis d'accroître fortement la transversalité.

Cet apport de RM à l'entreprise est précisé et complété par le second records manager interviewé. Premièrement, le RM contribue à l'efficacité opérationnelle de l'entreprise. Du point de vue économique, il se préoccupe de dépenser moins d'argent dans les processus de la gestion de l'information. Il contribue à la protection de l'entreprise dans son aspect sécuritaire : il conserve et détruit ce qui doit l'être. Au sein de l'organisation, le RM possède à la fois le rôle défensif et offensif : défensif pour protéger l'entreprise (aspect juridique) ; offensif dans l'amélioration de l'efficacité de l'entreprise (gains financiers). La dématérialisation en constitue un exemple. La transformation numérique de l'entreprise est aujourd'hui un axe fort impliquant le RM. C'est ainsi, que les principaux sujets en cours sont : le travail collaboratif, les réseaux sociaux de l'entreprise et les moteurs de recherche. En effet, l'information y est considérée dans une chaîne temporelle « *de bout en bout* » où l'archivage (conservation, patrimoine) est un des processus RM, les deux fonctions ne pouvant être séparées « *du point de vue opérationnel* ».

Les rôles endossés par le records manager de l'entreprise sont multiples : il formule des orientations stratégiques : propose la vision et les axes de travail. C'est dans ce rôle que le records manager formalise le schéma directeur de la gestion et de la conservation de l'information de l'entreprise à cinq ans. Il se

préoccupe de l'opérationnel et des aspects budgétaires ; il gère des équipes et des compétences liées. Ses rôles pourraient être réunis au sein de trois grandes fonctions : fonction managériale, fonction de pilotage et de l'expertise (RM, archivage, archivage historique, GED, dématérialisation, signature électronique, etc.). Etant à la fois au service des collaborateurs et des clients de l'entreprise, le RM est considéré comme fonction transversale de l'organisation.

Expérimentalement, l'entreprise s'est dotée de la politique RM et de l'archivage en 2006 et a formalisé le schéma directeur en 2016. Ce dernier énonce la stratégie de la gestion et de la conservation de l'information de l'entreprise à l'horizon de cinq ans. La continuité, l'adaptation, l'innovation constante, la réflexivité sous tendant l'historicité seraient les principales caractéristiques observées du RM sein de cette organisation.

Le second exemple de l'application est celui de la Bibliothèque nationale de France (BnF). Nous y avons rencontré deux records manager successifs (en 2008 et en 2009). Cette institution culturelle communique et partage largement (BnF, 2013) son importante expérience en matière de RM. La Bibliothèque nationale de France compte 2800 agents dont 1200 rédacteurs potentiels. La production documentaire journalière se situe entre 300 à 400 documents, plus de 800 documents par jour sont consultés par l'ensemble du personnel. Dans ce contexte, la direction de l'institution a souhaité optimiser la gestion documentaire pour éviter tout dysfonctionnement lié à la perte ou à la mauvaise circulation de l'information mais également « faire travailler ensemble » les acteurs grâce à une base de production commune et aux mêmes règles de gestion documentaire. Dans la conception du nouveau système, le records manager s'est inspiré des méthodes et techniques RM pratiquées à la Banque d'investissement européen au Luxembourg mais également du ministère de la justice, du ministère des affaires étrangères et de plusieurs sociétés privées.

Dans son retour d'expérience, le records manager a noté l'importance des deux éléments clés : la conduite du changement et la mise en place des moyens nécessaires afin d'assurer la réussite du projet. Le records manager a souligné également le rôle de premier ordre joué par les correspondants RM au sein de chaque unité dans la mise en place du système RM mais surtout dans sa continuité. Ils veillent à la bonne

application des procédures RM et la bonne utilisation du logiciel. Ils restent également à l'écoute des usagers. Ils sont identifiés comme source de proposition dans l'évolution des systèmes d'information et de la gestion documentaire. Mais, « *le pari n'est définitivement gagné que lorsque plus personne ne se souvient qu'il ait pu exister un autre mode de gestion auparavant. A partir de ce moment il est également possible de considérer qu'on peut passer à une conservation quasi exclusive de la forme électronique des documents* » (C. Dhérent). La BnF bénéficie aujourd'hui de douze années d'expérience dans le domaine RM.

Notre troisième exemple est un témoignage d'un archiviste, conservateur du patrimoine, concernant, tout particulièrement, la transformation numérique des institutions étatiques. Il fait constat d'une présence persistante d'une vision passéiste de l'archivage et de la difficulté pour le producteur d'archives de concevoir que l'histoire se fait à partir de l'instant présent. L'acteur ne fait pas toujours le lien entre le document qu'il produit et le document qu'il demande par la suite au service d'archives. Cela apparaît d'autant plus difficile dans le cas des documents numériques. Souvent, la vision d'ensemble de l'utilisateur sur le court, moyen et long terme de la vie du document n'est ni acquise ni comprise. Ce manque de partage d'une vision commune peut engendrer une rupture entre l'acteur, producteur des documents et le service qui les accueille. Cette situation renforce l'incompréhension des processus de gestion des documents. Pour palier à cette problématique, les archivistes forment des correspondants au sein des organismes où les compétences RM semblent rares voire absentes. Pourtant, ces compétences permettent de gérer les documents courants et intermédiaires et, ainsi, préparer les documents à l'archivage historique.

Du point de vue législatif « *la bascule dans le numérique [des services étatiques] est faite, nous avons changé d'univers* ». L'exemple constitue la récente loi pour une République numérique adoptée le 26 janvier 2016 (Assemblée nationale, 2016). Cependant l'archiviste note la difficulté d'acceptation de ce passage. La réflexion sur les problématiques de l'archivage numérique est en cours mais les pratiques liées à celui-ci ne sont pas encore acquises. L'usage du papier reste toujours très présent et l'archivage numérique est confondu avec la numérisation et le stockage des documents électroniques. De ce

fait, l'archivage électronique est souvent envisagé comme une problématique technique sans tenir compte du besoin de la conduite du changement des pratiques RM et/ou archivistiques et, en particulier, celles liées à l'environnement numérique.

L'automatisation des processus d'archivage numérique peine. Plusieurs problématiques sont à l'origine de cette situation. Les systèmes informatiques où on ne peut pas aller « du bout en bout » dans l'élaboration et le traitement du document (manque de signature électronique par exemple) engendrent les impressions à différents moment de processus. Cela impacte directement la gestion des dossiers d'affaire qui deviennent hybrides (papier et numérique, sur plusieurs systèmes, plates-formes) sans que le lien ne soit fait entre ses différentes composantes. Les moyens techniques manquent également. Les services archivistiques ne disposent pas toujours des SAE (système d'archivage électronique) nécessaires et le projet étatique VITAM (Premier ministre, 2015) ne dévoile pas facilement son contour malgré les travaux bien amorcés. Dans ce contexte, l'enjeu dans l'environnement numérique, est de penser l'archivage dès la conception du document mais également d'avoir un vocabulaire unique et unifié qui est celui de la gestion de l'information tout au long de son cycle de vie.

A travers ces trois exemples, nous pouvons lire l'importance de la contribution du RM et de l'archivage à l'efficacité et au bon fonctionnement des organisations. Ils démontrent le besoin d'une démarche RM envisagée et installée dans la durée. Ils font état du passage effectif vers le numérique et des problématiques qui accompagnent ce changement du point de vue des professionnels et des acteurs-producteurs de l'information. Le rôle premier de l'acteur (acteur-professionnel, acteur-producteur-usager, acteur-contributeur) dans le système RM est mis également en exergue.

RM et GSI : approches disciplinaires et professionnelles

D'ores et déjà en 1986 Marchand et Horton *in* (Bergeron, Dufour, Maurel, & Mercier, La gestion stratégique de l'information, 2010), proposent une approche de développement de la gestion stratégique de l'information où la gestion des archives courantes est une première phase de cette démarche en cinq étapes. Les phases suivantes sont : la gestion des technologies de l'information, la gestion des ressources

d'information commerciales, la gestion de la veille stratégique pour, enfin, atteindre le niveau de la gestion stratégique de l'information.

Selon la définition que l'on donne à la gestion de l'information (*information management*), ce domaine de recherche partage un certain nombre de paradigmes communs avec le RM comme, par exemple, la collecte, la diffusion et le traitement des documents. L'une des principales différences identifiée (Lee, 2015) entre ces deux disciplines est que le RM se focalise sur la capture, la gestion et la préservation de la valeur probante des documents.

La discipline étroitement citée en lien avec le RM est l'archivistique. Duranti (1989) soutient que les records managers et les archivistes représentent deux branches d'un seul métier qui partagent le même fond théorique (Lee, 2015, p. 16). Selon Sanford (Sanford, 2010), à partir des années 1970 dans les pays anglo-saxons l'archivistique et le RM deviennent deux champs disciplinaires différents avec les formations qui leur sont dédiées. Il y a été reconnu que les archivistes se consacrent au passé alors que les records managers s'intéressent au présent dans le but d'améliorer l'efficacité des organisations (Sanford, 2010).

Une autre discipline mentionnée en lien avec le RM est la gestion des connaissances, plus particulièrement dans son volet de la capitalisation de la mémoire-objet de l'organisation. Aussi, les travaux de Gunnlaugsdottir (2006, 2008) montrent que le système RM vient en support de la gestion et de l'organisation des connaissances organisationnelles.

Quelques textes professionnels tentent d'identifier la place de records management dans la démarche ingérée de la gestion de l'information au sein des organisations ainsi que les liens qu'il entretient avec d'autres méthodes et techniques qui en font partie. Nous présentons ci-dessous deux approches qui nous semblent illustratives : le document de normalisation : X 50 – 185 (Sutter, 2005) et le cadre de management de l'information du gouvernement de l'Australie occidentale.

Selon Sutter (Sutter, 2005), le schéma du management de l'information couvre quatre fonctions complémentaires de la gestion de l'information :

- 1) la gestion documentaire : focalisée sur l'information formelle d'origine externe et mise à disposition d'une manière neutre ;
- 2) la veille stratégique et l'intelligence économique, focalisées sur l'information externe informelle et traitée de façon orientée ;
- 3) le RM qui englobe une grande partie de l'information interne et externe, traitée comme preuve de l'activité et, enfin,
- 4) la gestion des connaissances qui englobe l'information interne (produite ou appropriée) comme support des savoirs et des savoir-faire, incluant l'information informelle et qui est associée à un processus de création de la valeur. Le texte n'analyse pas en profondeur les interdépendances de ces quatre fonctions, elles y apparaissent, notamment, à travers la cartographie normative (gestion documentaire : NF EN 82045-1, XP X 50-453 ; RM : ISO 15489; veille, sécurité et information stratégique : X 50-53, BP Z 74-500 ; gestion des connaissances et capitalisation de l'expérience : HB 275-2001, FD X 50-190). L'activité transverse indiquée par le fascicule est la gestion des documents vitaux (RM). Dans le contexte de l'analyse des risques, elle concerne tout type de l'information ainsi que des données.

Un document du gouvernement australien présente une approche innovante de la gestion stratégique de l'information tout en intégrant le RM (Government of Western Australia, Department of Local Government, 2013), (Michel, 2013). Ce document donne les définitions des sept « éléments clés » de « la discipline » de « la gestion de l'information » qui sont : la gestion des connaissances, la gouvernance, la sécurité, la gestion des actifs informationnels (*Information Asset Management*), l'accès et l'usage de l'information, la tenue des dossiers (*Record Keeping*) et enfin la gestion des données. Le RM interviendrait principalement dans le volet de « tenue des dossiers ». En effet, le « *record keeping* » désigne les activités qui permettent d'assurer que les exigences législatives et réglementaires soient respectées dans le traitement de l'information. Dans ce modèle, le RM n'apparaît pas dans la gestion des connaissances. Cela résulte de l'acception donnée à cette composante de la gestion de l'information définie comme la pratique consistant à extraire une valeur supplémentaire de l'information comprenant l'analyse et le *reporting* (Government of Western Australia, Department of Local Government, 2013, p. 18).

Les modèles cités montrent que le RM est bien identifié comme constituante indissociable des approches ingérées de la gestion de l'information dans les organisations. En revanche, les liens entre le RM et les autres composantes semblent moins explorés.

Conclusions

L'ère numérique bouleverse les théories sous-jacentes du RM et de l'archivage dont il partage bien des fondements. Les pratiques RM dans cet environnement numérique sont également en mutation et demandent de l'accompagnement. Les défaillances dans la gestion de l'information et des documents peuvent entraîner la prise de décision erronée et des pertes financières de l'établissement (Melinat, Kreuzkam, & Stamer, 2014). Alors, il s'opère aujourd'hui un changement organisationnel mettant en avant, notamment, la complémentarité des fonctions et des dispositifs de gestion de l'information. Cette démarche de regroupement se traduit, parmi d'autres, par la création de fonctions transverses de gestion informationnelle au sein des organisations (Bergeron, Dufour, Maurel, & Mercier, 2010).

A travers cet exposé, nous avons pu considérer les rôles du RM dans la gestion de l'information des organisations tant étatiques que privées. Le RM contribue à la construction et la préservation de la mémoire vive et vivante de l'organisation. De ce fait, il assure la continuité de son fonctionnement, il contribue aux gains économiques et à la sécurité de celle-ci. En ce sens, le records management constitue un des piliers de la gestion stratégique de l'information. Au sein de celle-ci, il entretient des liens étroits avec l'archivage historique, la gestion des connaissances et l'intelligence économique dans son volet sécuritaire, notamment.

Nos travaux futurs s'attacheront à explorer et préciser ces liens dans la démarche de la GSI.

Bibliographie

AAF. (2007). *Abrégé d'archivistique : principes et pratiques du métier d'archiviste*. Paris: AAF.

AAF, & ADBS. (2005). *Comprendre et pratiquer le records management : analyse de la norme 15489 au regard des pratiques archivistiques françaises*. Paris: ADBS.

AFNOR. (2004, Mars). FD X50-185 Mars 2004 Outils de management - Management de l'information. Saint Denis La Plaine.

AFNOR. (2008, Avril 2). Records management : gestion et conservation des documents en toute sécurité. Saint-Denis La Plaine .

Asemblée nationale. (2016, Janvier 26). Projet de loi pour une République numérique : adopté par l'Assemblée nationale en première lecture. Consulté le Avril 4, 2016, sur Assemblée nationale: <http://www.assemblee-nationale.fr/14/ta/ta0663.asp>

Bergeron, P., Dufour, C., Maurel, D., & Mercier, D. (2010). La gestion stratégique de l'information. Dans J.-M. Salaün, & C. Arsenault, Introduction aux sciences de l'information (pp. 183-205). Paris: La Découverte.

BnF. (2013, Novembre 22). La gestion des documents administratifs électroniques. Consulté le Avril 2, 2016, sur Bibliothèque nationale de France: http://www.bnf.fr/fr/professionnels/innovation_numerique/a.records_management.html

Cottin, M. (2015, Avril). Révision de la norme ISO 15489. Faire évoluer nos savoir-faire. I2D – Information, données & documents , 53, pp. 7-8.

Davenport, T. H., & Prusak, L. (1997). Information ecology : mastering the information and knowledge environment. New York: Oxford University Press.

Dhérent, C., Sibille de Grimouard, C., Ott, F., & Brouland, M. (2009, Février). Comprendre les enjeux et les fondamentaux (dossier records management). Documentaliste- Sciences de l'Information , 46, pp. 32-42.

Drouhe, T. G., Keslassy, G., & Morineau, E. (2000). Records management : mode d'emploi. Paris: ADBS.

Government of Western Australia, Department of Local Government. (2013). Information and Communications Technology Strategic Framework.

Gunnlaugsdottir, J. (2008). Registering and searching for records in electronic records management systems. International Journal of Information Management , 28, pp. 293–304.

Gunnlaugsdottir, J. (2006). The Implementation and Use of ERMS : A Study in Icelandic Organizations. Thesis. University of Tampere.

Heylighen, F. (2002). Complexity and Information Overload in Society: why increasing efficiency leads to decreasing control. *The Information Society* , pp. 1-44.

Joseph, P., Debowski, S., & Goldschmidt, P. (2012). Paradigm shifts in recordkeeping responsibilities: implications for ISO 15489's implementation. *Records Management Journal* , 22 (1), pp. 57-75.

Lee, S. H. (2015). A Bibliometric Analysis of the Field of Records Management: A Case Study of Records Management Journal, 1989-2013. Master thesis. Faculty of Information University of Toronto.

Lezon Rivière, A. (2013). Gestion stratégique de l'information et construction de sens : étude des pratiques informationnelles des chefs militaires. Thèse de doctorat. Université Paris 8 Vincennes-Saint-Denis.

Maceviciute, E., & Wilson, T. (2002). The development of the information management research area. *Information Research* , 7 (3).

Madsen, D. (2013). Disciplinary perspectives on information management : the 2nd International Conference on Integrated Information. *Procedia - Social and Behavioral Sciences* (73), pp. 534–537.

Melinat, P., Kreuzkam, T., & Stamer, D. (2014). Information Overload: A Systematic Literature Review. *Business Informatics Research* , pp. 72-86.

Michel, K. (2013, Septemebre 30). Définir un cadre de gestion de l'information, qu'est-ce que ça veut dire ? Consulté le Avril 2, 2016, sur FIL D'ARIANE: <https://karinmichel.wordpress.com/2013/09/30/definir-un-cadre-de-gestion-de-linformation-quest-ce-que-ca-veut-dire/>

Nowé Hedvall, K. (2007). Tensions and Contradictions in Information Management : an Activity-theoretical Approach to Information Activities in a Swedish Youth/Peace Organisation. Doctoral thesis. Swedish School of Library and Information Science, University College of Borås/Göteborg University.

Premier ministre. (2015, Mars 15). Vitam : vers un socle d'archivage électronique commun à toute l'administration . Consulté le Avril 4, 2016, sur Le portail de la modernisation de l'action publique:

<http://www.modernisation.gouv.fr/ladministration-change-avec-le-numerique/par-son-systeme-dinformation/vitam-vers-un-socle-d-archivage-electronique-commun-toute-l-administration>

Sanford, J. (2010, June). Capturing Records Management : A Holistic Analysis. Consulté le Avril 2, 2016, sur ARMA International Educational Foundation:
http://www.armaedfoundation.org/pdfs/Jamie_Stanford_Paper.pdf

Sutter, E. (2005). Le management de l'information : présentation commentée du document de normalisation X 50-185. Paris: ADBS.

Wilson, T. D. (2002). Information management. Consulté le Octobre 06, 2013, sur InformationR.net:
http://informationr.net/tdw/publ/papers/encyclopedia_entry.html

Services de gestion de données de recherche en Suisse : Vers une collaboration élargie & une offre exhaustive

Pierre-Yves Burgi

Université de Genève, (STIC)

pierre-yves.burgi@unige.ch

Eliane Blumer

Université de Genève, (STIC)

eliane.blumer@unige.ch

Basma Makhoulf-Shabou

Haute école de gestion de Genève (HEG)

basma.makhoulf-shabou@hesge.ch

Lydie Echernier

Haute École spécialisée de Suisse occidentale (HES-SO)

lydie.echernier@hes-so.ch

Résumé

Dans cet article, les auteurs décrivent l'étude de la gestion du cycle de vie des données en cours en Suisse avec un accent sur la préservation à long terme. Basé sur une grande analyse de documents, ainsi que des entretiens semi-structurés, le projet décrit ci-dessous vise à fournir des services au niveau suisse pour les besoins les plus importants dans la gestion des données

tout au long de leur cycle de vie, y compris un point d'accès et de contact, de formation et de conseil unique, des solutions de gestion des données actives, des options de stockage et de conservation à long terme, ainsi qu'une gouvernance pour tous les aspects.

Mots-clés: Services de gestion de données de recherche, la préservation des données numériques, la gestion des données de cycle de vie inclusifs.

Introduction

Les développements technologiques en cours dans le contexte numérique donnent lieu à de nouvelles façons de recueillir, capturer, stocker, manipuler et transmettre de gros volumes de données, et de stimuler la communication et la collaboration entre les scientifiques. Les données de recherche, une classe importante de contenu numérique, est définie par l'OCDE (2007) :

“factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings”.

L'information contenue dans les données de recherche constitue donc des atouts précieux pour les chercheurs. Le National Science Board des États-Unis (NSB, 2005) distingue trois types de données : observationnelles, de calcul, ou expérimentales.

Cette distinction est cruciale dans les choix faits pour l'archivage et la préservation de l'information. Dans les cas où les observations ne seront pas répétées (par exemple en astronomie, sciences de la terre, etc.), les données sont par conséquent uniques et incalculables. À l'inverse, pour les données de calcul, et si des informations complètes sur le modèle sont disponibles (généralement contenues dans les métadonnées), la conservation dans un dépôt à long terme n'est pas nécessaire, parce que les données peuvent être reproduites. Entre les deux, les données d'expériences, qui peuvent être reproduites, ne doivent pas être stockées indéfiniment ; mais en pratique, il peut ne pas être possible de reproduire précisément toutes les conditions expérimentales, et/ou les coûts de reproduction de l'expérience peuvent être prohibitifs. En dehors de ces types de données, d'autres peuvent être regroupées sous le terme de « record »

(Borgman, 2015) couvrant la documentation générale (du gouvernement, des entreprises, etc.), à laquelle on pourrait ajouter des cahiers de laboratoire électroniques dans le contexte de la recherche.

Dans la plupart des cas, les chercheurs ignorent l'importance du cycle de vie de leurs données, un concept qui englobe de nombreux aspects, dont certains pourraient dépendre de la discipline (comme illustrée précédemment). D'une manière générale, le cycle de vie des données commence par l'acquisition des données brutes. Par la suite, ces données sont analysées pour tester des hypothèses et valider des modèles. Les résultats de ces études conduisent généralement à des publications, une étape importante dans la vie de ces données traitées, car cela peut correspondre à la fin du projet et/ou de son financement. A ce stade, les données sont souvent négligées, généralement abandonnées sur un stockage acheté et autogéré par le chercheur, des serveurs improvisés et/ou des serveurs institutionnels (dans le meilleur des cas), et sont rapidement oubliées en négligeant l'opportunité de les promouvoir dans d'autres contextes. Des disciplines avec d'énormes quantités de données, telles que les sciences de la vie, dépendent de leur infrastructure institutionnelle ou externe respective. Le savoir-faire de la gestion des données est généralement transféré entre les équipes de recherche et ne provient pas d'un service extérieur. Même si les sauvegardes sont conseillées et utilisés, la perte des données constitue une question sérieuse et une problématique récurrente, indépendante de la discipline de recherche.

Les sections suivantes exposent la problématique étudiée (section 2) ; les principales considérations méthodologiques (section 3) ; l'analyse des besoins concernant les lignes directrices et les outils de gestion de données de recherche (GDR) en milieu universitaire suisse, effectuée dans le cadre du projet DLCM (section 4) ; la présentation du projet DLCM ; et des dimensions connexes en mettant l'accent sur la préservation à long terme (section 5).

Problématique

La préservation de l'information numérique est aujourd'hui assez bien maîtrisée (Schumacher et al. 2014), même s'il reste difficile de garder les flux bit inchangés sur de longues périodes de temps (Rosenthal, 2010). En effet, l'information est soumise

à la seconde loi de la thermodynamique, qui est que le désordre (ou entropie) tend à augmenter de façon permanente, de sorte que toutes les données sont nécessairement corrompues avec le temps. Cela implique que, pour préserver l'information, ce trouble naturel doit être écarté, par exemple grâce à la capacité de calculs, qui permet restaurer une information corrompue. Par contre, ce mécanisme a des limites : par exemple l'énergie qu'il requiert. Par conséquent, une sélection d'information à conserver devient une nécessité. Il est en tout cas établi que la croissance exponentielle des données de production, la multiplication des utilisations de l'ordinateur, l'obsolescence des objets, etc. augmentera la demande d'énergie (Bihoux, 2014).

Bien que la conservation soit la clé pour enregistrer les données de recherche, elle est par elle-même inutile si les différentes étapes du cycle de vie des données sont omises. Les chercheurs devraient en effet rendre explicite, à un stade précoce du projet, ce qu'ils ont l'intention de faire avec leurs données au cours du projet et, ensuite, s'ils envisagent de les partager (ou les garder confidentielles), de combien de temps ils ont besoin pour les préserver, etc. Ils devraient également documenter leurs ensembles de données afin d'être en mesure de les comprendre à l'avenir, pour eux-mêmes mais aussi pour d'autres chercheurs, et principalement au sein de la même discipline, mais pas seulement (Goodman et al., 2014; Wallis et al., 2013; Wallis, 2014). Ces plans de gestion des données (DMP) sont de plus en plus demandés par les organismes de financement et les institutions de recherche.

Une partie importante des formats de données concerne les DMP et leur obsolescence éventuelle. En principe, avant d'ingérer des données dans un référentiel, les chercheurs doivent se conformer aux formats recommandés afin que les cycles de migration puissent être assurés sur le long terme. Un tel respect n'est cependant pas toujours possible pour diverses raisons, qui comprennent la performance et la volumétrie, et aussi tout simplement la volonté des chercheurs à le faire. Dans de tels cas, la description des données, un processus long et complexe auquel les chercheurs se conforment rarement, précisément parce qu'ils préfèrent utiliser leur temps pour les tâches qu'ils jugent plus pertinentes pour faire avancer leurs recherches travail, devrait être développé de manière à devenir auto-décrit. Pour soutenir la préservation de toutes les données indépendamment du format, Klindt et Amrhein (2015) définissent deux niveaux de conservation : passive et active. Au niveau passif, les

données sont conservées seulement au niveau du bit, fournissant aux chercheurs la possibilité de stocker des données à un coût minimal, tout en respectant les contraintes d'éditeurs et/ou de bailleurs de fonds. De toute évidence, cela correspond à une demande de terrain. A l'inverse, au niveau actif, tous les mécanismes de préservation nécessaires sont appliqués pour s'assurer que les données restent interprétables tout au long des cycles de migration, mais cela nécessite un travail préparatoire plus intensif des producteurs de données. Dans les deux cas, les données conservées doivent rester aussi accessible que possible pour d'autres usages, et non conservés dans des archives sombres (dark archive). L'accès devrait par conséquent être facilité, ouvert autant que les formats standards le permettent, dans le but de servir d'autres chercheurs dans un futur proche et lointain. Des évaluations périodiques de valeur doivent être effectuées, de telle sorte que les coûts pour préserver les flux binaires correspondent à la valeur intrinsèque des ensembles de données conservées, étant donné que cette valeur peut diminuer avec le temps.

Combien de chercheurs ont à l'esprit toutes ces questions de gestion lors de l'acquisition de leurs (de plus en plus) grands volumes de données ? La préservation de l'information numérique est clairement un processus complexe et coûteux, mais qui ne peut plus être contourné. Les chercheurs ont besoin d'être au moins rendus conscients de ces problèmes grâce à la formation dispensée par des professionnels de l'information, et dans certains cas, aidés par des Data Scientists afin de maximiser la curation de leurs données. Le projet suisse que nous décrivons dans cet article a cette ambition de couvrir tous les principaux problèmes contenus dans la gestion du cycle de vie des données (Data LifeCycle Management – DLCM), sachant que la préservation de l'information numérique représente l'élément central de cette démarche ambitieuse.

Considérations méthodologiques

Cette section présente l'approche méthodologique générale, les principales techniques de collecte et d'analyse des données et un aperçu des principaux résultats attendus à la fin de la recherche, qui est encore en cours. L'étude actuelle a pour but de proposer des solutions méthodologiques et technologiques qui pourraient assurer une gestion du cycle de vie des données de recherche appropriée afin de garantir l'interopérabilité et l'exploitabilité.

Compte tenu de la nature exploratoire de cette étude, la réalisation de la recherche de DLCM est basée sur une approche qualitative. En ce qui concerne la collection de données et d'analyse, 3 étapes ont été réalisées :

- Une grande analyse de documents visant principalement à effectuer une revue de la littérature académique et professionnelle exhaustive avec laquelle le projet DLCM peut construire un ensemble de bonnes pratiques nationales et internationales en matière de politiques et de plan de gestion de données (DMP).
- Des entretiens semi-structurés pour confirmer les besoins du chercheur ; environ 50 entretiens semi-structurés ont été réalisés avec des chercheurs de différents départements dans une variété de disciplines (voir la section 4), offrant une connaissance approfondie des diverses pratiques de données de recherche dans les universités suisses.
- L'analyse des données recueillies à la fois de la revue de la littérature et des entretiens semi-structurés, pour aboutir à un aperçu des principales considérations qui devraient être prises en compte pour offrir un cadre/modèle général pour la gestion du cycle de vie des données de recherche rationnelle, avec les lignes directrices connexes, des outils, des compétences et des technologies pour permettre sa mise en œuvre effective dans le milieu universitaire suisse.

Résultats attendus

À la fin de l'étude de DLCM, les résultats suivants sont attendus :

- Un plan de gestion des données (DMP) adapté aux communautés de recherche suisses
- Des lignes directrices pour les chercheurs et les professionnels de l'information sur la gestion des données de recherche (y compris des directives pour les données et la diffusion de métadonnées)
- Un modèle de stratégie pour les établissements d'enseignement supérieur
- Un portail suisse sur les données de recherche (avec des outils et des pratiques recommandées)
- Les rapports sur l'analyse des lacunes et des concepts pour une infrastructure de stockage distribué à l'échelle

nationale et conforme à la norme OAIS (Open Archival Information System ou Système ouvert d'archivage d'information)

- Une boîte à outils pour la construction SIP – AIP – DIP (paquet d'information à verser – paquet d'information archivé – paquet d'information diffusé) à partir des sous-ensembles de données de recherche (y compris l'interface graphique adaptée aux différents outils)
- Un prototype d'une infrastructure OAIS conforme et évolutive
- Des modèles d'affaires pour la prestation de services de conservation à long terme, y compris pour la formation et le conseil
- Un inventaire des modules de formation DCLM existants (y compris les réseaux d'experts en collaboration avec d'autres projets suisses)
- Des modules spécifiques de formation DCLM et des modules d'enseignement pour l'intégration dans des cursus de sciences de l'information
- Une participation à des ateliers, des conférences nationales et internationales

Validité de la recherche

Pour assurer la qualité de notre étude, l'équipe du projet DLCM investit dans plusieurs activités. En ce qui concerne la validité scientifique :

- Un état de l'art en ce qui concerne le travail effectué dans ce domaine avec des chercheurs locaux, nationaux et internationaux ;
- Les principales théories, des normes, des projets sont examinés et pris en compte dans le domaine de la gouvernance des données de recherche ; et
- Les résultats intermédiaires sont partagés et publiés dans des conférences internationales avec des pairs examinateurs.

En ce qui concerne la validité pratique, l'équipe de DLCM a rencontré plusieurs professionnels dans les institutions publiques et les entreprises privées qui ont partagé leur retour d'expérience dans le domaine de la conservation des données à long terme.

Ajouté à cela, le caractère interdisciplinaire de cette étude (informatique, curation de données, sciences de l'information,

etc.) offrira un impact important au projet. Enfin, pour assurer une collaboration fluide entre de multiples partenaires et des chercheurs de différentes disciplines, un glossaire terminologique a été défini et partagé lors des réunions régulières, des ateliers, etc.

Analyse des besoins

Comme mentionné précédemment, une analyse des besoins exploratoire a été menée afin d'obtenir une connaissance plus approfondie des besoins des chercheurs et des solutions en place. Pour cela, chaque institution partenaire a conduit des entretiens semi-structurés pendant deux mois (septembre et octobre 2014).

La structure de ces entretiens comportait quatre parties principales, à savoir: (1) les données initiales et flux de travail, (2) l'analyse et les données d'exploration, (3) les publications, l'archivage et la gestion des données à long terme, et (4) l'avenir : les défis, risques et perspectives. Le tableau 1 présente la compilation de toutes les disciplines interrogées.

Institution	Nombre d'entretiens	Disciplines
UNIGE	8	Théologie, informatique, linguistique, allemand, neurosciences cognitives, sciences de l'éducation, de la géomatique, archéologie, vulnérabilité, sciences politiques, médecine (recherche sur le cancer des enfants), sciences politiques
ETHZ	8	Biotechnologie et génie biologique, réseaux sismiques, sociologie, comportement du consommateur, optique quantique, informatique scientifique/recherche avec des photons, physique
UNIL	15	Médecine sociale, sciences sociales, humanités numériques, génomique, biologie des systèmes, bio-informatique, santé publique, Imaging and Media Lab, recherche sur le cancer
EPFL	5	Transport et mobilité, optoélectronique quantique, nanomatériaux

		supramoléculaires et interfaces, laboratoire de communication audiovisuelle, virologie et génétique
UNIBAS	7	Recherche en biologie (Biozentrum, 2), biologie (installations de base, 2), psychologie moléculaire, santé publique (STPH), humanités numériques
UNIZH	5	Sciences juridiques, biologie/microscopie, biologie/protéomique, hôpital universitaire, géosciences
Total	49	

Tableau 1 : Résumé des entretiens

Comme le montre le tableau 1, les disciplines interrogées sont diverses. Dans une prochaine étape, chaque entretien a été inscrit dans un tableau récapitulatif, organisé par discipline et expédié dans les quatre parties principales de l'entretien avec, le cas échéant, une classification plus fine en fonction des similitudes dans les réponses (voir Annexe). Sur cette base, une analyse des principales tendances a été réalisée. Cette analyse montre que la gestion des données dépend évidemment des stratégies institutionnelles et/ou des habitudes de recherche dans les disciplines spécifiques considérées. La plupart des disciplines interrogées sont confrontées à divers défis au cours des différentes étapes de leurs projets de recherche. En fait, le cycle de recherche d'aujourd'hui dépend de solutions ad-hoc ainsi que d'habitudes spécifiques au sein des départements de recherche.

Les données initiales : traitement de workflow

Tout d'abord, peu ou pas de DMPs formels sont utilisés, à moins que les instances de financement en exigent un au moment de la candidature du projet. En conséquence, la perte de données et la description sont souvent mentionnées comme des difficultés.

Un autre défi, qui concerne la description des données et de stockage, est qu'il n'y a pas de lignes directrices communes entre les disciplines, et donc les données existent dans une pluralité de formats (vecteur, vidéo, audio, image, texte,

graphique, flux binaires brutes, et ainsi de suite), propriétaires et/ou ouverts, en fonction de l'application logicielle. Ces formats sont adaptés aux besoins des projets de recherche (et de l'équipe) et rarement dans l'optique de la conservation des données.

Même si des normes de description communes existent dans certaines disciplines internationalement bien organisées, comme en géographie, en sciences humaines et en sciences de l'éducation, ici, les normes ne sont pas utilisées, avec même, dans certains cas, la réponse à la question de la signification exacte d'une «donnée» qui peut étonnamment s'avérer difficile.

En ce qui concerne le stockage des données, dans la plupart des cas, des serveurs improvisés et autofinancés sont utilisés, étant donné que les départements informatiques institutionnels sont souvent plus lents dans la fourniture de solutions que le taux des données produites par les chercheurs. Indépendamment de la discipline de recherche, les chercheurs sont conscients de la nécessité de sauvegarder leurs données, la perte de données étant une préoccupation reconnue. Cependant, l'organisation des back-ups n'est pas toujours considérée comme la tâche exclusive de l'institution, mais aussi de l'individu.

Analyse et données d'exploration

Le plus grand défi dans cette partie semble être la liberté au sein de l'organisation des données. Chaque projet de recherche, département, parfois même chercheur a son/ses propres habitudes qu'il est difficile d'harmoniser. En ce qui concerne le partage et la préservation, à l'heure actuelle, Dropbox reste une solution utile, même si elle est soumise à la loi américaine.

Publication, archivage et gestion des données à long terme

La notion de conservation à long terme est généralement absente et s'avère être une difficulté majeure. Combien de temps avons-nous besoin de conserver les données ? Quelle est la meilleure stratégie pour la préservation à long terme ? Existe-t-il des lignes directrices ou des règles existantes ? Ce sont des questions que les chercheurs se sont posées et pour lesquelles ils ne disposent pas encore de réponse.

Une autre difficulté est le droit d'auteur, car il n'y a pas de vision claire de qui est le propriétaire des données spécifiques

ou de ce qui doit être fait afin de publier des données sans violer le droit d'auteur.

Les données de recherche à l'avenir : défis, risques et perspectives

Même si les opinions relatives à l'importance de la conservation à long terme en ce qui concerne les dépôts diffèrent, un point est régulièrement mentionné : les chercheurs indiquent qu'il n'y a pas de réponse adéquate à la question de ce qui pourrait et devrait se produire pour les données de recherche après la fin d'un projet et/ou la publication réussie des résultats scientifiques. Souvent, les données disparaissent dans le désert de bureaux ou des serveurs plus ou moins bien entretenus. Alors que la survie de ces données est une préoccupation, leur bonne gestion est citée comme une question plus difficile.

Les chercheurs interrogés ont répété l'importance d'une incitation au partage des données. Une telle incitation pourrait être, par exemple, la citation de données de pairs. Pour ce faire, des ensembles de données doivent être identifiés de façon permanente par un DOI (Digital object identifier, littéralement « identifiant numérique d'objet »). Avoir une personne supplémentaire pour gérer leurs données serait apprécié, ainsi que la possibilité d'obtenir plus de citations.

Les disciplines avec une énorme quantité de données voient aussi un risque relatif aux coûts croissants des lieux de stockage, avec la question connexe de savoir qui va payer pour cela.

Les besoins

Sur la base des défis mentionnés ci-dessus, les besoins suivants ont été identifiés pour le projet DLCM :

- Des lignes directrices et du soutien pour aider les chercheurs à gérer et à stocker leurs données, des solutions informatiques et d'analyse ad hoc
- Des solutions pour le stockage des progrès de la recherche (gestion de données actives) avec des instantanés périodiques de jeux de données définis
- Le développement et/ou la maintenance de référentiels de données de recherche en ligne à plus long terme

Le projet DLCM

Cette section présente tout d'abord un aperçu des grands axes qui composent cette étude et la seconde partie met l'accent sur la préservation à long terme.

Axes du projet DLCM

Le projet DLCM est organisé en 5 axes principaux :

- 1) les lignes directrices et politiques ;
- 2) les données actives de recherche ;
- 3) la conservation à long terme ;
- 4) le conseil, la formation et l'enseignement ; et
- 5) la diffusion.

Les sections suivantes présenteront brièvement les objectifs de chacun de ces axes.

Lignes directrices et politiques DLCM

Cet axe vise à définir des lignes directrices fondées sur une analyse exhaustive de pratiques académiques et professionnelles et sur une revue de la littérature avec laquelle le projet DLCM peut construire un ensemble de bonnes pratiques nationales et internationales en matière de plan de gestion de données (DMP) et de politiques applicables dans les établissements d'enseignement supérieur.

Données actives de recherche

Cet axe vise à fournir aux chercheurs un large éventail de solutions technologiques et des bonnes pratiques pour gérer leurs données actives, avec un accent particulier sur la collecte, le traitement et l'analyse de ces données.

Quatre objectifs principaux ont été définis :

- Des cahiers de laboratoire électronique (ELN), des systèmes de gestion de l'information de laboratoire solutions (LIMS) et un soutien pour les sciences de la vie,
- Virtual Environment Research (VRE) pour les humanités digitales,
- Des solutions de travail pour les (grandes) installations scientifiques,
- Des solutions de stockage de logiciel pour les données actives.

Conservation à long terme

Cet axe vise à établir un pont entre les données actives et la conservation à long terme, et veut fournir des solutions de publication. Pour ce faire, nous allons nous baser sur des concepts bien établis, tels que le Curation Life Cycle et le modèle OAIS.

Cinq parties principales sont également prévues :

- Un Gap Analysis des dépôts actuellement utilisés dans les projets partenaires,
- Une boîte à outils pour construire des paquets d'information OAIS (SIP-AIP-DIP),
- Un prototype d'infrastructure OAIS conforme et évolutive,
- Une interface graphique (GUI) adaptée aux différents outils et apte à interfacier une infrastructure OAIS conforme,
- Un modèle d'affaires pour la préservation à long terme des données.

Conseil, formation et enseignement

Cet axe vise à mettre en place de services de conseil, de formation et d'enseignement dans le domaine de DLCM. Pour ce faire, nous entendons :

- Fournir des services de formation DLCM avec les besoins généraux et spécifiques,
- Créer des services de conseil DLCM au sein de chaque institution partenaire, qui sera coordonné par un bureau central, le cas échéant,
- Adapter les connaissances acquises aux cursus de Bachelor et de Master,
- Développer des cours en sciences de l'information, où les élèves seront formés et pratiqueront directement les connaissances acquise, assurant ainsi la pérennité de ces connaissances pour les générations futures.

Diffusion

Cet axe vise à promouvoir les résultats du projet DLCM et à établir des contacts et des collaborations avec des institutions et d'autres projets pertinents qui ne sont pas directement impliqués ou liés au projet DLCM. Pour cela, plusieurs campagnes de sensibilisation sont prévues au cours du projet.

Préservation

La conservation des données de recherche sur le long terme est l'un des objectifs principaux du projet du fait qu'il sert à établir un pont entre la gestion de données active et le moyen de stocker des données (sur le long terme) afin de donner accès à l'ensemble de données qui a abouti à une publication, ou simplement de permettre leur réutilisation pour faire avancer la recherche.

Gap-analysis

Vu que le projet DLCM regroupe huit institutions, certaines d'entre elles disposent de solutions pour la conservation à long terme, mais elles sont dans la plupart des cas principalement utilisées pour les publications d'archive. Nous avons commencé le projet avec une analyse des lacunes afin d'évaluer les différents stades de maturité et de conformité avec l'OAIS. En outre, un large panel de technologies a été rencontré au sein de ces institutions : l'Université de Zurich (UNIZH) exploite un référentiel basé sur le logiciel Eprints. L'ETH à Zurich (ETHZ), utilise Rosetta d'Ex Libris, un référentiel basé sur Invenio, avec l'intention de le lier à Zenodo à l'avenir (à Lausanne - EPFL), tandis que les Universités de Genève et Lausanne utilisent Fedora Commons pour leurs dépôts. Pour effectuer cette analyse, nous avons appliqué les outils d'évaluation développés par l'équipe POWRR (Préservation des objets numériques avec des ressources restreintes). La méthodologie définit cinq domaines fonctionnels (le stockage et l'emplacement géographique, la fixité des fichiers et l'intégrité des données, la sécurité de l'information, les métadonnées et les formats de fichiers) et quatre niveaux de conservation numérique (protection, savoir, surveillance, et réparation des données). Il en résulte une grille d'évaluation (figure 1), qui représente l'intersection du cycle de vie de la conservation numérique du Centre curation numérique et de l'OAIS Reference Model (ISO 14721:2012). D'après cette grille, il est évident qu'aucune institution ne remplit complètement la conformité requise, même si certaines (par exemple l'ETHZ) sont proches de la cible. En outre, la plupart des dépôts ne sont pas conformes à la « préservation de Bit fiable à long terme » caractéristique, qui est une condition nécessaire pour la conservation à long terme, mais un défi (voir ci-dessus). En effet, il est reconnu que le stockage de disque reposant uniquement sur RAID, les systèmes ne détectent pas tous les types d'erreurs et sont soumis à « la

corruption silencieuse des données » (Rosenthal, 2010; Bairavasundaram et al., 2008), ce qui nécessite donc d'autres mécanismes de checksum de niveau supérieur pour garantir que les erreurs soient détectées.

	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Geographically Dispersed Data Storage Model	Exit Strategy	Migration/Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost	
ETH Zurich	(X)	X	X		X	X	X	X	X	X	X	(X)	X	X	X	X	(X)	X	X	X		X	(X)
UNGE Archive/Fedora3	X				X	X	X	X	X	X	X	X	X	X	X	X	X				X	X	Free
UNL Server/Fedora3	X	(x)	X	X	X	X	X	X	X	X	X	X	X	X	X	(x)	X	X	(x)		X	X	Free
UDN/ZORA Eprints	(X)		X		X	X	X	X	X	X	X			(X)	X	X					X	X	Free
Zenodo (EPFL)	(X)	X	X		X			X	X	X	(X)	X	(X)	(X)	X	X		X	X	X	X	X	Free

Figure 1 Grille d'évaluation. (X) indique les similarités

Boîte à outils pour construire des paquets d'information OAIS (SIP-AIP-DIP)

À un certain stade de leur travail de recherche, les chercheurs souhaitent sélectionner un ensemble de données à partir du stockage de données actif pour l'ingérer dans un dépôt de stockage à long terme. Les motivations des chercheurs pour accomplir cette étape (passant de l'actif à l'état semi-actif ou passif) sont diverses et principalement les suivantes : les éditeurs demandant un accès durable aux données utilisées pour obtenir les résultats dans le document publié, les chercheurs ont besoin d'un DOI (ou tout autre identifiant permanent) pour partager ouvertement l'ensemble de données, ou l'archivage des données simplement à la fin d'un projet de recherche. Pour que cette étape soit aussi souple et transparente que possible, les chercheurs devraient avoir la possibilité de pousser leurs données sélectionnées « d'un clic » à partir de la zone de stockage actif dans un référentiel de données (état semi-actif) et/ou dans un système de conservation à long terme (état passif). Ensuite, le modèle de référence OAIS nécessite la préparation d'une trousse d'information Soumission (SIP). Dans le processus, plusieurs micro-services sont généralement activés et

comprennent les calculs de somme de contrôle, détection de format, vérification de virus, etc., avec, à la fin de l'ensemble une archive Information Package (AIP). Il existe des outils pour fournir de tels micro-services automatisant les AIPs (par exemple Archivematica, Workbench conservateur, iRODS, etc.). Du côté de la consommation, les interfaces utilisateurs spécifiques doivent être conçues pour permettre aux utilisateurs de récupérer des informations à partir du référentiel et/ou à partir de l'AIP archivé. Dans ce dernier cas, cela se fera en générant une trousse d'information Diffusion (DIP) délivrée à l'utilisateur qui en a besoin.

Infrastructure mesurable et conforme à l'OAIS

Une autre contrainte importante du modèle de référence OAIS concerne le stockage physique. Le stockage doit être hautement redondant, auto-correcteur, résilient et doit être composé de multi-copies réparties géographiquement, tout en maintenant l'intégrité et la traçabilité de l'information stockée. Safe PLN est un projet en cours pour la construction d'un tel OAIS conforme, sur la base de l'infrastructure LOCKSS (Maniatis et al. 2005) et regroupant plusieurs partenaires en Europe et au Canada . LOCKSS est basé sur le concept de la tolérance aux pannes byzantine, initialement connu comme le problème des généraux byzantins (Lamport et al., 1982). Pour être en mesure de se défendre contre les défaillances byzantines, dans lesquelles les composants d'un système échouent avec des symptômes qui empêchent certains composants du système de parvenir à un accord entre eux, la redondance est nécessaire. En outre, pour tolérer deux défaillances de nœuds, au moins 7 nœuds sont nécessaires. Dans la version actuelle, seulement 200 Go peuvent être archivés sur 7 nœuds géographiquement éloignés. En outre, et même si dans le futur Safe PLN pourrait être étendu à la gestion de volumes beaucoup plus importants, typiquement plusieurs pétaoctets, ou plusieurs téraoctets, nous considérons l'utilisation d'une infrastructure semblable à celle du Zuse Institute de Berlin (ZIB) (Klindt et Amrhein, 2015), construite sur Archivematica, pour la partie entrées, iRODS pour le stockage d'archives (principalement pour la gestion de deux réplifications, l'une étant géographiquement éloignée) et Fedora Commons pour exposer le DIP à travers une interface utilisateur dédiée. Parmi les nombreuses questions encore à résoudre pour rendre cette infrastructure évolutive, l'une est technique : comment vérifier et reproduire des pétaoctets de données

réparties sur des centaines de millions de fichiers ? Une autre question est liée au financement de l'infrastructure sur le long terme, et un modèle d'affaires est en cours d'élaboration selon la méthode décrite dans la section suivante.

Viabilité

Comme l'étude de DLCM est limitée dans le temps et les membres du projet sont hétérogènes, une stratégie pour la viabilité des services devait être mise en œuvre. Pour cela, le modèle d'affaires « Business Model Canvas » et sa proposition de valeur « Value Proposition Canvas », conçus par Alexander Osterwalder et Yves Pigneur et résumant la façon dont un établissement fait, fournit et gagne la valeur, ont été choisis.

Proposition de valeur et Business Model Canevas

Dans la proposition de valeur (voir la figure 2), on décrit généralement son/sa client/e sur le côté droit, y compris son/ses emplois, ainsi que les gains et les douleurs. Dans une autre étape, on tente de trouver des créateurs de gains et de douleurs et de les convertir en produits et services. Ainsi, chaque service créé répond à un besoin existant et préalablement identifié, et fournit une valeur viable.

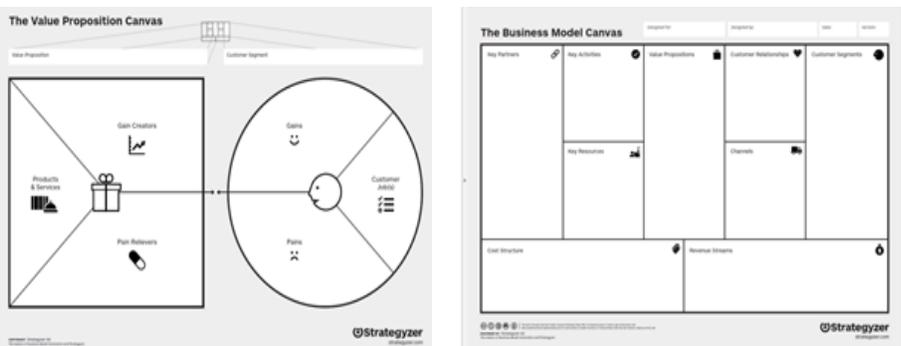
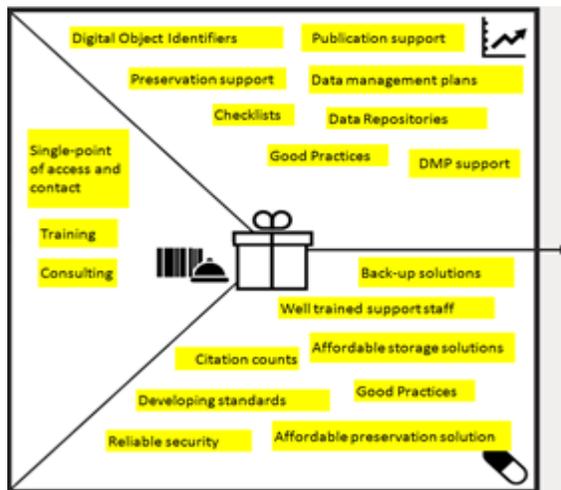


Figure 2: Value Proposition Canvas (gauche) and Business Model Canvas (droite)

Dans l'étape suivante, on reprend la proposition de valeur et on l'intègre dans un modèle d'affaires, alors que la description du client passe dans le champ « Segment client » et la proposition de valeur dans le champ correspondant. Ensuite, nous sommes en mesure de remplir les autres parties du modèle afin de décrire et de finaliser les services ciblés.

Dans le cadre du projet DCLM, la préservation à long terme est l'un des domaines de services qui seront couverts. Les deux prochaines images présentées à la figure 3 illustrent la façon dont le développement de services pour une préservation à long terme pourrait ressembler. Il est à noter que les étiquettes sont les premières hypothèses fondées sur les disciplines interrogées et leurs défis connexes (voir annexe). Les principaux services pourraient être : un point d'accès et de contact, une formation unique et des conseils en ce qui concerne la conservation à long terme.



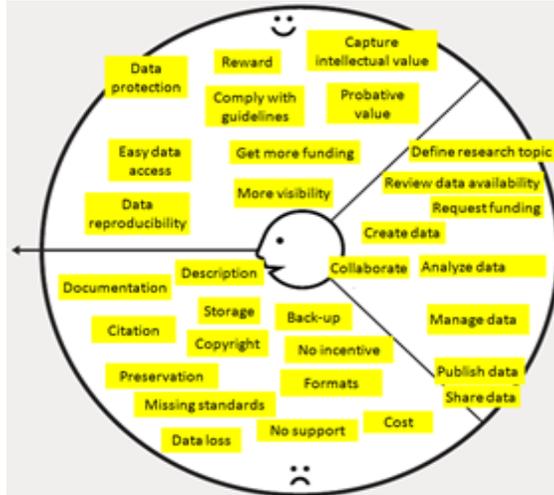


Figure 3: Value Proposition for the LTP service

Modèles de coûts

Un élément clé du modèle d'affaires est la structure des coûts, qui devrait correspondre aux flux de revenus. La question des coûts de la conservation à long terme a été abordée par un certain nombre d'études internationales, notamment le KRDS (Keeping Research Data Safe) projet (Beagrie, 2008 ; Beagrie, 2010). Ce projet, portant sur des études de cas au Royaume-Uni, tire des conclusions sur les principaux facteurs de coûts et, sur cette base, propose un cadre d'établissement des coûts contenant deux éléments majeurs : des ajustements économiques et de services. Les ajustements économiques se composent de l'inflation (salaire par exemple), la déflation (par exemple des supports de stockage), l'amortissement des actifs (coûts historiques/achat à travers sa vie utile) et les frais de retour sur investissement. Les ajustements de service comprennent les coûts liés au (1) personnel, (2) les phases d'acquisition, de cession et d'intégration, (3) le stockage d'archives, la planification de la conservation et la gestion des données. Fait intéressant, la sauvegarde et le stockage de fichiers à long terme ne représentent qu'une infime partie des coûts totaux. Selon cette étude, l'acquisition et le nombre d'entrées représentent jusqu'à 42%, l'accès 35%, et le stockage 23%. Encore plus extrême est une typologie des activités considérant la création de données, qui compte pour 73%, contre la curation et le stockage, représentant respectivement 24% et 3%. En dehors de l'étude KRDS, un autre travail remarquable dans ce domaine est l'étude

comparative de plusieurs modèles de coûts réalisée dans le cadre du projet européen 4C.

Parmi tous ces travaux, un modèle en particulier a retenu notre attention : le coût total de la préservation (TCP) de l'UC Curation Center (Abrams et al. 2012), qui englobe les coûts économiques associés à la préservation à long terme des actifs numériques. Ce modèle prend en compte 11 activités de préservation (système, services, serveurs, personnel, producteurs, Workflows, types de contenu, de stockage, de surveillance, d'interventions), et considère deux modèles de prix : « Pay-as-you-go » et « Payé ». Si, au premier abord, il revient à considérer un modèle complet des coûts, bénéfique pour la détermination de la charge de financement, en pratique, il apporte plus de questions que de réponses. En effet, pour être applicable, on doit d'abord savoir à l'avance le nombre de « producteurs ». Et comme le nombre de producteurs augmente, les coûts associés diminuent, ce qui signifie qu'un producteur a intérêt à venir tard pour bénéficier de prix avantageux. Deuxièmement, l'évaluation de toutes les activités de conservation est loin d'être triviale et/ou fiable. Par conséquent, nous avons développé un modèle de coût simplifié pour un TCP indépendant du nombre des producteurs, sur la base du « modèle de prix pay-up » sans assumer aucun retour sur investissement, mais qui prend en compte un coût global pour le maintien de l'infrastructure, y compris les allocations de personnel et un pourcentage annuel de baisse des prix pour le matériel. Nous sommes conscients qu'il existe d'autres coûts importants liés à des services tels que la création et la conservation des ensembles de données. Cependant, nous voulons être en mesure d'évaluer les coûts de base associés au stockage en toute sécurité de l'information, en supposant que d'autres services pourraient être facturés séparément en fonction du niveau de services que les clients demandent. Et pour les chercheurs, avoir des coûts raisonnables pour stocker leurs données sur le long terme est clairement une incitation importante pour éviter de laisser des données finir leur vie sur une infrastructure de stockage inappropriée.

Prochaines étapes

L'étude de DLCM a débuté en Suisse en Septembre 2015 et est toujours en cours. Par conséquent, les prochaines étapes sont les suivantes :

- Créer des modèles de viabilité pour tous les axes du projet mentionnés au chapitre 5.1,
- Les valider et de les adapter en interrogeant les utilisateurs,
- Travailler de manière itérative vers les résultats attendus et les besoins exprimés,
- Partager et collaborer au sein des institutions partenaires et au-delà de leurs limites institutionnelles afin de faire avancer le sujet.

Conclusion

Dans cet article, les auteurs décrivent l'étude de la gestion du cycle de vie des données en cours en Suisse avec un accent sur la préservation à long terme. Basé sur une grande analyse de documents, ainsi que des entretiens semi-structurés, le projet décrit ci-dessus vise à fournir des services au niveau suisse pour les besoins les plus importants dans la gestion des données tout au long de leur cycle de vie, y compris un point d'accès et de contact, de formation et de conseil unique, des solutions de gestion des données actives, des options de stockage et de conservation à long terme, ainsi qu'une gouvernance pour tous les aspects.

Les résultats ont montré que les chercheurs sont confrontés à plusieurs défis en matière de gestion de données, mais ont tendance à utiliser des solutions ad-hoc. Les principaux besoins exprimés sont des lignes directrices et un soutien dans la gestion des données, des solutions de stockage ad-hoc, des solutions informatiques et d'analyse, des solutions pour le stockage des progrès de la recherche et des référentiels. A la fin du projet, les services veulent répondre à ces besoins à l'échelle suisse et aller au-delà. En mettant en œuvre une stratégie de viabilité, à savoir le Business Model Canvas, les membres du projet DLCM utilisent le même vocabulaire et collaborent de manière itérative au-delà de leurs limites institutionnelles.

En fait, l'expérience de ces derniers mois a montré qu'il y a aussi plusieurs défis à relever, aussi bien au niveau d'un tel projet qu'au niveau de la réalisation de ses services à l'échelle nationale. L'orientation suisse à l'échelle du projet, ainsi que les différentes natures et besoins des institutions partenaires, demandent une bonne coordination et une gouvernance qui comprennent également la liste de tous les aspects stratégiques,

la gestion des risques tels que l'optimisation des résultats, ainsi que leur impact.

L'avenir proche montrera précisément comment les services et les résultats du projet seront décrits. Dans tous les cas, il est nécessaire de créer des synergies et des pratiques au sein de toutes les institutions suisses, qui peuvent être mutualisés.

Bibliographie

Abrams S, Cruse P, Kunze J, Mundrane M (2012) Total Cost of Preservation (TCP): Cost Modeling for Sustainable Services. UC Curation Center, California Digital Library.

Bairavasundaram LN, Goodson GR, Schroeder B, Arpaci-Dusseau AC, Arpaci-Dusseau RH (2008) An analysis of data corruption in the storage stack. FAST'08: 6th USENIX Conf. on File and Storage Technologies, 223–238

Beagrie N, Chruszcz J, Lavoie B (2008) “Keeping research safe. A cost model and guidance for UK universities” Final Report, Charles Beagrie Limited

Beagrie N, Lavoie B, Woollard M (2010) “Keeping research safe 2” Final Report, Charles Beagrie Limited

Bihouix P (2014) L'Âge des low-tech. Paris, Seuil.

Borgman CL (2015) Big data, little data, no data. MIT Press, Cambridge MA.

Goodman A, Pepe A, Blocker AW, Borgman CL, Cranmer K, et al. (2014) Ten Simple Rules for the Care and Feeding of Scientific Data. PLoS Comput Biol 10(4). doi:10.1371/journal.pcbi.1003542

Klindt M, Amrhein K (2015) One core preservation system for all your data. No exception! Proc. of the 12th International Conf. on Preservation of Digital Objects, 101-108

Lampert L, Shostak R, Pease M (1982) “The Byzantine generals problem” ACM Trans. on Programming Language and Systems 4(3) 382-401

Maniatis P, Roussopoulos M, Giuli TJ, Rosenthal DSH, Baker M (2005) The LOCKSS Peer-to-Peer Digital Preservation System. *ACM Transactions on Computer Systems*, 23(1) 2–50

NSB (2005) Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century. NSB-05-40, National Science Foundation.
<http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf>

OCDE (2007) OECD Principles and Guidelines for Access to Research Data from Public Funding. HYPERLINK
"<http://www.oecd.org/science/sci-tech/38500813.pdf>"
<http://www.oecd.org/science/sci-tech/38500813.pdf>

Osterwalder & Pigneur, *Business Model Generation* (2011).

Rosenthal DSH (2010) Bit Preservation: A Solved Problem? *The International Journal of Digital Curation*, 5(1) 134–148

Schumacher J et al (2014) From Theory to Action: “Good Enough” Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions. A Digital POWRR White Paper for the Institute of Museum and Library Services

Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, et al. (2011) Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6(6).

doi:10.1371/journal.pone.0021101

Wallis JC, Rolando E, Borgman CL (2013) If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE* 8(7).
doi:10.1371/journal.pone.0067332

Wallis JC (2014) Data Producers Courting Data Reusers: Two Cases from Modeling Communities. *International Journal of Digital Curation*, 9(1) 98–109

Annexe : résumé des entretiens semi-structurés orienté sur les défis des personnes interrogées

Disciplines	Initial data and workflow	Analysis and data exploration	Publication, archiving and long-term management	Research data in the future: challenges, risks, perspectives
Theology	No DMP Back-up	Freedom of methodology to organize data	Freedom of organization	No local storage Copyright issues
Informatics	No DMP	Lack of description	Lack of description	No storage for sharing and versioning Incentive
Linguistics	No DMP Description rules Back-up Lack of documentation	Freedom of methodology to organize data	Conservation rules Data is obsolete	Format issues
German	No DMP Description rules Back-up	Freedom of methodology to organize data	Conservation rules	Citation rules No archive Incentive
Cognitive Neuroscience	DMP for US-projects Data loss Local back-up and storage	-	Conservation rules No reproducibility	Versioning Standardization
Educational Sciences	No DMP Local storage and back-up	Freedom of methodology to organize data	Copyright issues	Institutional storage on local level with access rules Standardization
Geomatics	No DMP	-	-	Need of a data archive

Disciplines	Initial data and workflow	Analysis and data exploration	Publication, archiving and long-term management	Research data in the future: challenges, risks, perspectives
Archeology	Description rules Data storage No DMP	Freedom of methodology to organize data	Conservation rules Copyright issues	Interdisciplinarity
Vulnerability	No DMP No description rules Local storage	-	Data loss Data publication	-
Political Sciences and International Relations	Naming Preserving	-	-	data repository Twice production of same data Incentive
Medicine (Child Cancer Research)				
Quantum Optoelectronics	Local storage			Link between data and publication Cost Guidelines Storage Incentive
Transport and Mobility Laboratory	Local backup	Dropbox	Conservation rules	Checklists Institutional repository Incentive Lack of documentation Copyright issues

Disciplines	Initial data and workflow	Analysis and data exploration	Publication, archiving and long-term management	Research data in the future: challenges, risks, perspectives
Supramolecular Nanomaterials and Interfaces Laboratory	No DMP Back-up	-	-	Standardization Metadata
Genetics	Standardization in description	Dropbox	-	Guidelines Reproducibility
Audiovisual communication laboratory	-	-	-	Cost National perspective Metadata Incentive
UZH	Missing standards	-	-	Reproducibility Best Practices Incentive
Unispital ZH	Missing standards	Ethics	No long-term preservation strategy	Cost Centralized storage with access management Incentive
Proteomics	-	Centralized sharing tool		DMP Storage cost Long-term preservation
Photon Science/Scientific Computing	-	-	-	Cost Incentive Politics

Disciplines	Initial data and workflow	Analysis and data exploration	Publication, archiving and long-term management	Research data in the future: challenges, risks, perspectives
Biosystems Science and Engineering	-	-	Long-term preservation Conservation rules	Easily accessible web front end for data publication Infrastructure
Seismic Networks	-	-	Sharing (others can see, but not process)	Automatic snapshots Storage facilities
Social Sciences, Modelling, Simulation	No DMP	-	-	Indexing large volumes of data including hidden annotations Framework for management support Capture defined states of data periodically Differential record of algorithms' versions Comprehensive Swiss cloud solution Real time streaming of big data Reliable citation counts Realtime data input Support and staff

Disciplines	Initial data and workflow	Analysis and data exploration	Publication, archiving and long-term management	Research data in the future: challenges, risks, perspectives
Consumer Behavior	No DMP Data loss Reproducibility	-	-	Managing data all in one Local storage Incentive
Quantum Optics Group	-	-	-	Periodical snapshots of specific data Persistent link to published data Time
Physics, IT Services Group	No DMP Local Storage	-	-	Formats Incentive
Biocenter	No DMP	-	-	Repository
Epidemiology	-	-	-	Interoperability Long term storage is not funded
Digital Humanities	Data loss	-	Copyright	-
UNIL				

Pratiques et enjeux de la documentation numérique personnelle

Nasreddine Bouhai

Laboratoire Paragraphe (EA 349)

Équipe Citu-Paragraphe

Université Paris 8

COMUE Université Paris-Lumières

Résumé : Dans cet article, nous nous intéressons particulièrement à la question de la maîtrise et de l'exploitation de l'information en général et la gestion documentaire personnelle et celle des références bibliographiques en particulier ; ainsi que les pratiques et les enjeux de la gestion de l'information à l'ère du numérique et de la diversité des supports : ordinateur, tablette, smartphone cloud, etc. Face à la masse très importante de documents en ligne, les différentes démarches de recherche d'informations aboutissent souvent par l'accumulation d'une masse d'informations qu'il faut ensuite gérer pour une utilisation ultérieure, dans des contextes divers : mémoire, article, thèse, etc. La gestion des résultats souvent sous forme liens (signets) ou téléchargement de documents, représente une part toujours plus importante lors d'une recherche documentaire et le besoin à recourir à des outils et des techniques de gestion documentaire est plus que nécessaire, indexation/classement, gestionnaires de marque-pages et de références bibliographiques.

Mots-clés : Gestion documentaire, document numérique, bibliothèque personnelle, documentation personnelle, normalisation, système documentaire, références bibliographique, recherche d'information, indexation, classement, visualisation, navigation, archive, optimisation.

1. Introduction

La masse considérable d'informations, de ressources, de documents très divers qui caractérise cet espace numérique qui est le Web et dans lequel est produite et diffusée une quantité très considérable de ressources et de documents ; L'utilisateur doit recourir à une stratégie de recherche et de gestion documentaire, à savoir, identifier et sélectionner l'information pertinente dans la masse disponible, juger la qualité et la fiabilité de cette information par l'élaboration de la recherche, l'analyse et l'évaluation de l'information retrouvée. Face à ce gigantesque réservoir d'information, les différentes démarches de recherche d'informations aboutissent souvent par l'accumulation d'une masse d'informations qu'il faut ensuite gérer pour une utilisation ultérieure, dans des contextes divers : mémoire, article, thèse, etc.

La gestion des résultats souvent sous forme de signets ou de références, représente une part toujours plus importante lors d'une recherche documentaire et le besoin à recourir à des outils et des techniques de gestion documentaire est plus que nécessaire, plus spécialement, les gestionnaires de marque-pages et de références bibliographiques. Il existe un ensemble d'outils en version local¹³ ou réseau¹⁴ (Zotero, Delicious, Diigo..), propriétaire ou libre, commercial ou gratuit. Ces outils spécialisés, permettent la gestion et l'organisation des marque-pages et des références sélectionnées durant la démarche de recherche documentaire et leurs utilisations ainsi que la possibilité de les partager. Les notions de partage et de collaboration associées à la première grande mutation du Web pour désigner le Web 2.0 en utilisant une multitude de services où chaque usager du Web peut devenir acteur de l'information. L'un des exemples qu'on peut donner et la mise en commun des favoris, qui est une source de classification collaborative pour les réseaux sociaux.

Avec le développement de la toile depuis quelques années, ces outils et services se sont multipliés et leurs usages. Beaucoup de travaux universitaires mettaient déjà en avant l'intérêt de ces approches de partage et d'échange centrés usager, comme HyWebMap [BOU 02], IronWeb[DUS 00], Nestore[BEL 99], HNLS [BOU 04] et KWeb Organiser[BOU 02] etc., avec point

¹³Desktop Application

¹⁴Web Application

commun de ses outils, l'agrégation des connaissances dans un cadre individuel ou collaboratif.

Dans cet article, nous nous intéressons particulièrement à la question de la maîtrise et de l'exploitation de l'information en général et la gestion documentaire personnelle et celle des références bibliographiques en particulier ; ainsi que les pratiques et les enjeux de la gestion de l'information à l'ère du numérique et de la diversité des supports : ordinateur, tablette, smartphone *cloud*, etc.

2. Le document, l'information et l'information documentaire

Il est important de définir dans le contexte de ce travail, les notions de « document » et de l'« information » avant de les aborder un peu plus loin. Le document se définit avant tout comme un support et une information : *« ensemble formé par un support et une information, généralement enregistré de façon permanente, et tel qu'il puisse être lu par l'homme ou la machine »* (définition ISO)¹⁵, et plus largement comme un *objet porteur d'information*. Document et information vont de pair, une information a toujours besoin d'un support et le support et l'information forment un document. Le support désigne la nature matérielle du document, comme le support papier et le support numérique. L'information se matérialise dans une médiation à caractère documentaire. Quant à la notion de « donnée », *elle renvoie à une « représentation d'une information sous une forme conventionnelle adaptée à son exploitation »*¹⁶, la frontière entre les deux notions, « information » et « donnée » est assez mince, la notion de « donnée » joue le rôle d'intermédiaire entre document et information.

3. La gestion documentaire à l'ère de multi-supports

L'avènement de l'ordinateur personnel et du réseau Web, ont été des éléments moteurs pour le domaine de l'innovation, de la communication et de la diffusion de l'information. Le développement du Web et du numérique ont participé à une surabondance d'informations disponibles en ligne et un nombre de documents impressionnant. La recherche d'informations se

¹⁵International Organization for Standardization (Organisation Internationale de Normalisation)

¹⁶<http://www.academie-francaise.fr/dictionnaire>

fait désormais par le biais d'outils de recherche en ligne, avec la multiplication des recherches et l'usage de l'ordinateur au quotidien, dans des contextes différents, y compris pédagogique et scientifique. L'ordinateur (le bureau, figure 1.) fait office d'espace de notes et de dépôt documentaire au fil du temps, cet espace devient rapidement surchargé et ingérable, avec des noms de documents impertinents pour pouvoir retrouver un contenu rapidement sans passer par une recherche pleine texte à chaque fois et une perte considérable de temps. À l'ère de la généralisation de l'informatique et des ordinateurs, beaucoup d'utilisateurs passent énormément de temps, à rechercher des informations, des notes, des documents, qu'ils soient reçus par la boîte de messagerie, ou rédigés et déposés sur localement sur ordinateur. Cette démarche très présente chez beaucoup d'utilisateurs, illustre le besoin et la nécessité à recourir à une stratégie de gestion documentaire.



Figure 1. Exemple de bureau Windows surchargé, Source Google images

3. Maîtrise et exploitation de l'information : une question de méthodologie

3.1. Contexte et enjeux

La publication scientifique repose essentiellement sur les notions de la recherche et d'écriture de textes que ce soit un rapport, un article ou une thèse, elle suppose surtout une recherche d'informations très approfondie dans le domaine et l'exploration de la littérature scientifique à ce sujet. Cette démarche scientifique doit prendre appuie sur de précédents travaux et aussi un besoin de confrontation aux énoncés concernant le sujet ou les sujets connexes et dont la finalité est de trouver, choisir et exploiter de l'information pour développer une réflexion personnelle. Une réflexion qui prendra sa maturité qu'avec le croisement des différentes informations glanées à travers les documents, retenus et analysés, cela donnera forcément de la crédibilité scientifique au travail de recherche et il convient de bien référencer les différentes sources de lecture sous forme de « références bibliographiques ».

Si le but d'une de recherche documentaire reste la finalisation rigoureuse d'un travail scientifique ou académique, cela ne peut se faire sans une recherche d'informations approfondie et une veille informationnelle (dynamique et automatisée) pour une richesse documentaire de qualité. Ainsi, l'organisation et le classement de ces documents et de ses références ainsi obtenus, relève sans doute de la mise en place d'une stratégie documentaire bien maîtrisée et suivie, l'un des points-clés pour réussir une production de qualité.

3.1. Recherche d'informations

La recherche documentaire exige la mise en place d'une méthode rigoureuse et efficace et la maîtrise des différentes étapes qui consistent avant tout, à repérer, récupérer et traiter des données publiées et disponibles en version numérique ou pas. Cette démarche est primordiale à toute synthèse des connaissances et revue de la littérature dans un domaine précis. Elle doit être la plus pertinente possible et tendre vers une possible exhaustivité. Elle nécessite :

la maîtrise des outils et des stratégies de recherche ;

la maîtrise de l'évaluation et de la qualité de l'information ;

la maîtrise de la gestion et l'organisation des documents et des références trouvées ;

la maîtrise des normes et des règles de formatage et de citation des références ;

la maîtrise des outils bibliographiques.

3.2. Veille informationnelle automatisée et dynamique

Avec le développement des infrastructures réseaux et les avancées dans le domaine des technologies de communication, le document électronique s'est avéré comme le médium privilégié pour la diffusion de l'information. Pour la période des années 2000, la majorité des journaux scientifiques proposaient déjà des versions électroniques de leurs publications [REN 09]. Le besoin croissant des utilisateurs à consulter des articles scientifiques [TEN 08], se traduit en général et en amont par une activité de recherche documentaire et aussi la nécessité de gérer le flux d'informations qui en découlait ; un suivi et une veille documentaire sont nécessaires pour faire évoluer la richesse de la base documentaire. La veille informationnelle est tout d'abord de faire venir l'information à son environnement numérique au lieu d'aller la chercher sur le réseau. Une tâche récurrente pour beaucoup d'utilisateurs, et qui demande à recourir à des outils informatiques pour son automatisation. Les outils de veille n'en manquent pas : Pearltrees, Delicious, Diigo, HNLs, etc.

3.3. Organisation et classement de la documentation

La problématique de la gestion de la documentation personnelle, qu'il s'agisse d'une collecte de données en provenance d'un environnement externe comme Internet (moteurs de recherche) ou non (productions personnelles), se pose continuellement quand aucune démarche d'organisation et de structuration n'est mise en place bien en amont des processus de recherche ou de production. La question de l'accès rapide à l'information ne peut se concevoir que dans la mesure où l'utilisateur a suivi un processus de structuration, de stockage de l'information ainsi glanée ou produite au fil des besoins, pour ne pas se retrouver submergé par des masses des documents inexploitable. Une démarche grandement facilitée par l'outil informatique, les méthodes d'indexation et de classification sont deux solutions largement utilisées dans ce contexte.

3.3.1 Les pratiques folksonomiques des réseaux sociaux

Une folksonomie est une indexation sémantique et un mécanisme de l'évolution du Web vers le Web 2.0, elle se définit comme un ensemble de tags fourni par les utilisateurs pour indexer des contenus sur le Web. Les utilisateurs choisissent librement les tags pour une meilleure description de documents (ou de contenus), une indexation plus pertinente et une recherche facilitée par des simples mots-clés. Il existe une panoplie d'outils qui proposent des fonctionnalités d'indexation sémantique intuitive, une sorte de classification collaborative qui tend à valider la fiabilité des ressources sélectionnées et proposées par les utilisateurs. Une démarche qui confère à ce type de structure une compréhension par les machines, en exploitant les mots-clés pour classer, voire suggérer de nouvelles recherches [BOU 06] pour un effet « Boule de neige » [SAL 01]. Les sites de « *social tagging* », comme *Del.icio.us*, *Diigo*, *Flicker*, sont fondés sur une pratique folksonomique et permettent une classification collaborative de ressources, par des communautés d'utilisateurs.

3.3.3. Les cartes heuristiques¹⁷ : une approche native folksonomique

Les cartes heuristiques sont des représentations graphiques d'idées, appelées aussi carte mentale, arbre à idées... Ces représentations souvent sous forme arborescente (figure 2.), permettent la structuration de thématiques exprimées en relations sémantiques entre concepts (idées). Le contenu se développe hiérarchiquement du général (les nœuds parents) au plus spécifique (les branches sous-jacentes). L'utilisation de ces cartes dans un contexte de gestion documentaire, permet aisément d'y associer des documents multimédia locaux ou en ligne, les annoter et de les enrichir par des mots-clés. Pour un usage individuel, cette rapproche offre un moyen efficace pour une organisation documentaire et une veille informationnelle ciblées ; Le partage de ces cartes met en évidence les deux principes collaboratif et social de ces outils, essentiellement pour ceux qui proposent des déclinaisons, sur ordinateur, tablette, smartphone, pour une accessibilité nomade.

¹⁷Les mind map

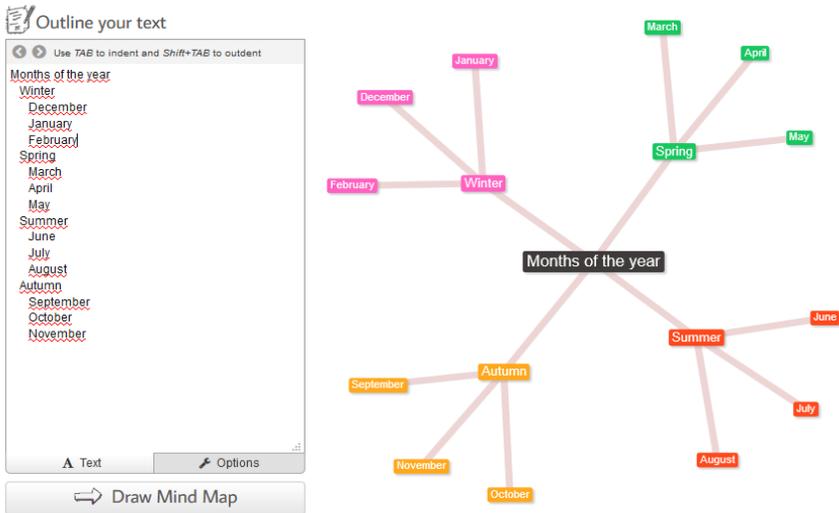


Figure 2. Carte heuristique sous text2mindmap¹⁸

3.3.4. Les outils intuitifs de partage des signets : *Pearltrees*

Dans la catégorie des outils intuitifs, *Pearltrees* est l'outil le plus abouti graphiquement, il propose une interface visuelle dynamique (figure 3a) et permet d'organiser et de partager des contenus (marque-pages) sous forme de carte heuristique(figure 3b). Il offre un ensemble de fonctionnalités utiles pour :

- la construction de cartes ;
- le partage des cartes entre utilisateurs ;
- le rapprochement sémantique des cartes par des connexions entre elles ;
- le travail de groupe ;
- l'exportation et l'intégration des cartes au sein des sites Web ;
- faire de la veille des marque-pages.

Pearltrees est un réseau social par un rapprochement des profils en utilisant surtout les centres d'intérêt des utilisateurs.

¹⁸<https://www.text2mindmap.com/>

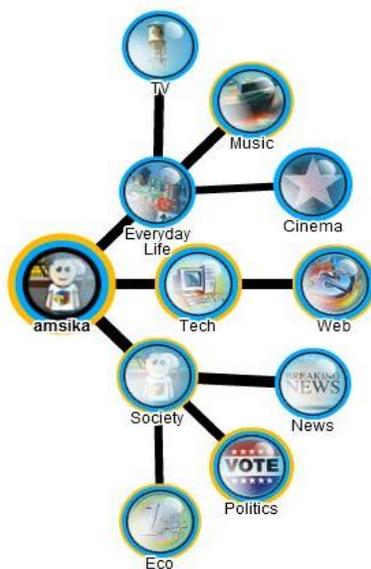


Figure 3a. Exemple de carte Pearltree ¹⁹

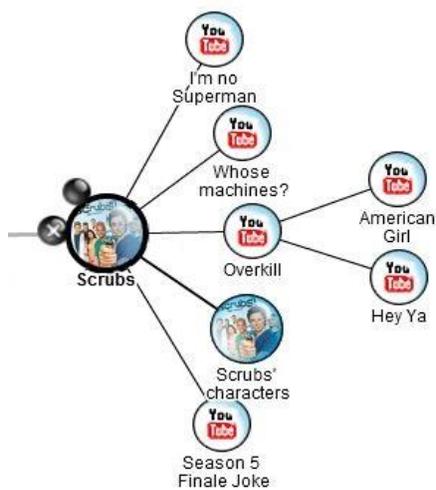


Figure 3b. Personnalisation de style des cartes²⁰

¹⁹<http://blog.pearltrees.com/>

²⁰<http://neptunya.fr/pearltrees-nep2013/>

4. La gestion des signets : une pratique documentaire renouvelée

Mémoriser la trace des consultations documentaires lors des différentes recherches en bibliothèque ou sur le Web devient une tâche récurrente tout au long d'une recherche documentaire. Cette trace est représentée par des références et des signets, deux éléments importants qu'il faut organiser et structurer pour un usage ultérieur, valoriser et argumenter une rédaction d'un côté, partager et en faire profiter la communauté de l'autre.

4.1. Les signets de navigateurs : des mécanismes rudimentaires !

Avec les navigateurs et le mécanisme de signets, le problème reste toujours posé quand il s'agit d'organiser et de structurer les références trouvées lors des différentes recherches. Le travail d'organisation doit être fait manuellement et un effort important doit être fourni pour obtenir une base de références correctement structurée et facile à exploiter. La difficulté s'accroît avec l'avènement du multi-support, l'utilisateur peut-être amené à changer d'outil de recherche, passant d'un ordinateur de bureau au portable, d'une tablette au smartphone et une dispersion de ses traces de recherche au fil du temps. Le problème de ce type de recherche nomade trouve sa solution dans les outils d'agrégation, de synchronisation de références et le *cloud computing* comme espace de sauvegarde documentaire accessible sans limite géographique.

Des systèmes précurseurs comme HyWebMap [BOU 02], IronWeb [DUS 00], Nestor [BEL 99] ont amélioré l'usage des favoris, en proposant des mécanismes de partage des favoris personnels. Les utilisateurs choisissant d'adopter IronWeb ou HyWebMap, mettent à disposition des autres utilisateurs l'ensemble de leurs favoris. Ce mode de partage réside dans ces collections de favoris « expertisés » qui forment alors des « autorités » dont l'usage en terme de recherche pourra s'avérer plus pertinent que les résultats produits par les moteurs de recherche généralistes.

4.2. Fédérer des ressources documentaires

Le Web a fait émerger la notion de marque-pages (signets, favoris) dès les premières montures des navigateurs, avec des simples mécanismes de mémorisation, qui permettent de

sauvegarder les liens de pages Web considérées comme pertinentes par l'utilisateur. Un classement rudimentaire pour rapprocher un ensemble de ressources dans un même dossier, une forme primaire d'organisation qui n'exprime en rien une éventuelle relation sémantique entre documents. L'utilisateur exprime ce rapprochement par des noms de dossiers et de rubriques évocateurs qui joueront plus le rôle d'une classe sémantique fédératrice des ressources ainsi trouvées. Les avantages et les inconvénients des mécanismes de marque-page n'en manquent, une base de marque-pages locale, non accessible à distance, cette difficulté est accentuée par le fait multi-supports ; la simplicité est l'avantage principal pour marquer des liens fréquemment utilisés.

4.3. Les outils Web 2.0

Devant la masse d'informations disponibles sur Internet et le nombre très important de ressources gratuitement accessibles ou des adresses des ressources de bibliothèque en ligne, la gestion des résultats obtenus à l'issue des multiples recherches d'informations, représente une part de travail toujours plus importante dans une démarche de recherche. L'usage des outils de gestion bibliographique reste le moyen adapté pour un gain de temps et d'efficacité. Les avancées technologiques laissent apparaître de nouveaux modes de gestion de l'information basés sur les notions de partage et d'échange des résultats de recherche d'information.

Contrairement à l'approche traditionnelle qui consiste à l'usage des « marque-pages » des navigateurs (Favoris sous IE, Marque-pages sous Firefox...), avec ses inconvénients et ses avantages et dont le besoin d'accéder souvent et rapidement à des pages fréquemment visitées reste indispensable ; Il en va différemment quand il s'agit de documents qui ne sont pas consultés régulièrement. Pour pallier ces difficultés, des outils en ligne ont vu le jour avec le Web 2.0 ; Et dont la démarche consiste à une gestion nomade des signets, des processus de travail collaboratif, partage et échange des références de ressources, ce qui lui confère un côté social. Les mécanismes des signets sociaux simplifient « la distribution des listes de référence, de bibliographies, papiers, et d'autres ressources entre pairs et étudiants »[DRE 09], et participent pleinement à la veille informationnelle nécessaire à beaucoup d'utilisateurs.

La définition des signets sociaux ²¹ est la « pratique qui permet de sauvegarder ses signets sur Internet, de leur attribuer des mots-clés pour en faciliter la recherche et de les partager avec une communauté d'utilisateurs. »²². Le principe général de ces services est de gérer les favoris (bookmarks) sur un serveur distant, contrairement aux navigateurs²³, ce qui leur confère un usage nomade et permettre le partage avec d'autres utilisateurs. Il existe plusieurs outils de ce genre actuellement sur le Web, tels que *Diigo* ou *Delicious*, et d'autres issus de certains travaux de recherche et précurseurs dans ce domaine comme *KWeb Organizer*²⁴.

4.3.1. Diigo

*Diigo*²⁵ est un service de gestion de signets en ligne, créé en 2006. Il s'inscrit dans la même lignée que le célèbre pionnier *Delicious*, le service *Diigo* propose aux utilisateurs de manière assez simple de gérer leurs signets en ligne, et de permettre à d'autres utilisateurs de les consulter et de les exploiter (pour permettre le partage de liens, et « la navigation sociale »). Cet outil a fait une percée technologique remarquable depuis 2009 en intégrant un ensemble de fonctionnalités très innovantes²⁶, ce qui lui confère un grand intérêt pour les différents usages, professionnel, individuel ou collaborative. Par rapport à la concurrence, *Diigo* propose certaines nouveautés comme :

croisement d'*URLs* des différents utilisateurs et la constitution de communautés dont les centres d'intérêt sont assez proches ;

une *toolbar*, pour un usage simplifié à partir des navigateurs et accéder à une grande partie des fonctionnalités du compte *Diigo* ;

une fonction d'archivage (copie en cache) ;

la notion de groupe thématique avec possibilité d'y associer des fils de discussion.

4.3.2. Delicious

²¹ En anglais *Social bookmarking*

²² Définition [Office québécois de la langue française, 2008]

²³ Firefox, IE, Opéra, Chrome...

²⁴ Outil issu des travaux du Laboratoire Paragraphe de l'université de Paris 8

²⁵ Acronyme pour *Digest of Internet Information, Groups and Other Stuff*

²⁶ L'intégration de la technologie Furl de Looksmart

Parmi les nombreux outils de gestion des signets en ligne, on trouve *delicious*, une référence incontournable de ces dernières années. La popularité l'outil ne se dément pas avec des millions d'utilisateurs et d'URLs en sa base de données. Il n'en reste que sa stagnation technologique lui a fait perdre ces derniers temps sa place de leader dans le domaine à la faveur de son concurrent *Diigo*. Le projet *delicious* a été créé fin 2003 par un certain Joshua Schachter dont l'objectif initial comme pour beaucoup de projets sur le Web est de répondre à un besoin personnel, sauvegarder ses propres signets et les rendre accessibles de partout.

Faire de la veille informationnelle est aussi possible que simple, le système permet de profiter de la recherche collaborative de ses membres et cela par une simple veille par indication des mots-clés à chercher. Avec son API *delicious*²⁷, le système offre beaucoup de possibilités de développement de petites applications Web : moteur de recherche, nuage de mots-clés, etc., en utilisant des flux JSON et surtout avec la généralisation de l'usage des flux RSS²⁸, il sera plus facile l'intégration de ces applications au sein des sites Web.

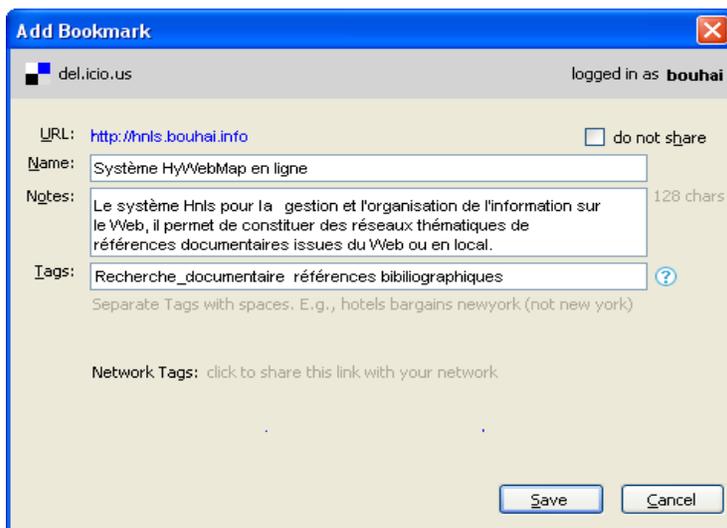


Figure 5. Opération d'ajout d'un signet avec Delicious

²⁷ Pour *Application Programming Interface*

²⁸ Il existe 3 formats XML utilisés pour la syndication de contenu Web

Le service Delicious propose une extension pour le navigateur web Mozilla Firefox pour rendre plus facile la gestion et l'accès à l'espace personnel ouvert au sein du service en ligne²⁹. Cette extension consiste à installer de nouveaux boutons sur la barre d'outils de Firefox, c'est-à-dire des raccourcis et permettre ainsi l'ajout et l'enregistrement des liens de pages visitées, directement au sein de l'espace personnel (compte) *Delicious*.

4.3.3 HNLS : HyWebMap On Line System

Le système HNLS est une transposition version réseau de l'outil HyWebMap pour un accès nomade et collaboratif, l'outil reprend l'ensemble des fonctionnalités de son prédécesseur avec moins de contraintes liées à l'usage des logiciels sur poste : difficultés d'installation, de mise à jour et de stockage de données... Le HNLS permet à un utilisateur de :

gérer ses recherches Web en organisant ses marque-pages sous la forme de cartes thématiques (document-réseau)³⁰ [BOU 05];

échanger et partager ses cartes avec d'autres utilisateurs ;

croiser les éléments communs de ses cartes avec d'autres et obtenir des méta-réseaux [BOU 05] ;

utiliser les cartes comme espace collaboratif de recherche d'informations ;

enrichir les documents associés aux cartes par des annotations et des mots-clés ;

faire de la veille automatisée et dynamique ;

Les spécificités du Web en font un environnement très difficile à exploiter. L'usage des annotations est une voie pour le partage et l'échange d'informations, autrement, atteindre l'objectif initial du Web en occurrence un environnement collaboratif où chaque usager est lecteur et auteur. Les intérêts des annotations pour l'usager sont multiples, de la compréhension et la relecture d'un document déjà annoté [MAR 98] à la personnalisation en fournissant une trace d'utilisation ainsi que le partage facile des informations. Les annotations témoignent essentiellement d'un engagement personnel du lecteur avec un document [MAR 98][CAD 00].

²⁹ delicious.com

³⁰Document-réseau ou document-réticulaire

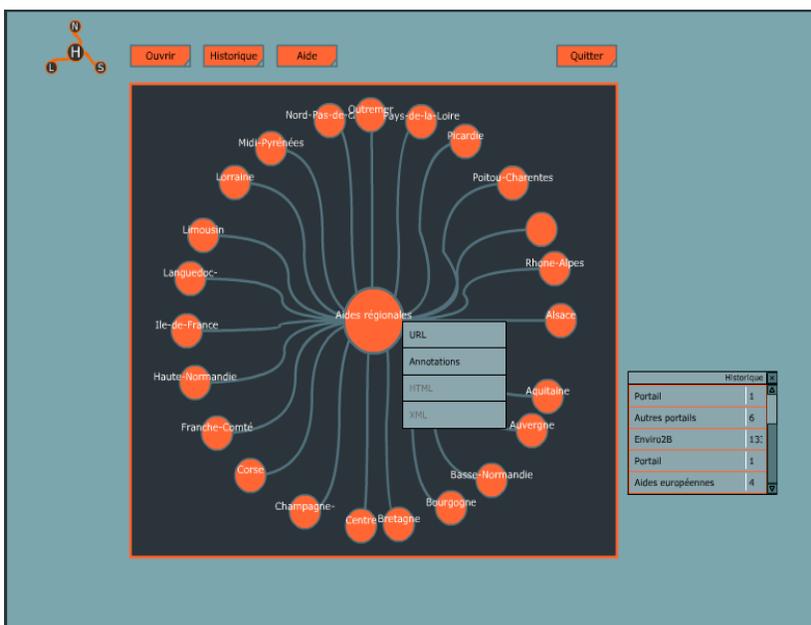


Figure 5. Capture de l'interface de HNLS

5. Gestion des références bibliographiques

Les citations donnent un aperçu de la densité de l'argumentation dans un travail de recherche scientifique et de l'état de l'art du domaine. L'utilisation des outils pour références bibliographiques remonte bien loin que l'apparition de l'Internet, certains outils en fonctionnement local existaient déjà pour répondre à des besoins bien précis de certaines communautés : chercheurs, étudiants... Dans une réciprocité innovation<->usage, les innovations en science de l'information et de la communication ont fait évoluer les usages dans ce domaine, ainsi, les besoins des utilisateurs ne cessent d'évoluer au fil du temps et influencent à leur tour le développement des outils. Les différents outils, anciens et récents tentent plus au mois à suivre les nouvelles pratiques, collaborative et sociale des utilisateurs, certains complètent leur tableau de fonctionnalités pour répondre aux nouvelles exigences des utilisateurs, d'autres stagnent et disparaissent : *Reference Manager...*

5.1. Les outils de gestion des références bibliographiques

Depuis les années 2000 les offres des logiciels libres pour la gestion documentaire généralement et à la gestion des références bibliographiques spécialement, ne cesse de s'enrichir. Beaucoup d'outils à degré différent soit au niveau richesse fonctionnelle ou au niveau qualité sont disponibles et évoluent continuellement. Ces outils attirent de plus en plus du monde, étudiants, enseignants et chercheurs, dont il existe un grand besoin dans la gestion des résultats des recherches documentaires relatives à leurs travaux universitaires. Cela n'exclut pas le monde professionnel et amateur de la lecture et de la culture, dont le besoin d'organiser et de gérer leurs références personnelles n'est plus à négliger. Ils permettent un gain de temps pour retrouver, comparer, regrouper, citer des références ou établir des listes bibliographiques selon des normes reconnues lors d'un travail de rédaction (mémoire, article, thèse, etc.). Ce type d'outils permet de :

- récupérer des notices bibliographiques depuis des bases de données ou des catalogues ;
- rentrer manuellement des références bibliographiques ;
- rappeler ses références à l'intérieur d'un document sous forme de notes ;
- constituer une bibliographie formatée à la fin d'un document ;
- partager ses références bibliographiques ;
- stocker des documents.

Parmi ses solutions libres pour la gestion des références bibliographiques, un nombre important reste dans la sphère du développement, des nouvelles versions apparaissent sur le marché avec des fonctionnalités pratiques liées au Web³¹. Cette évolution permet aux utilisateurs de trouver des solutions à leurs besoins même si aucun logiciel ne peut donner une satisfaction complète. La possibilité de coupler des solutions en utilisant des éléments intermédiaires communs³².

5.1.1. Zotero

³¹ Collaboration, version plugin, fils rss, etc.

³² Le format BibTEX et l'usage de XML est de plus en plus courant.

Zotero est un autre outil de gestion bibliographique assez présent sur le Web, il est initié par le *Center For history and New Media*³³ et disponible en open source. Il propose un nombre de fonctionnalités pour organiser, enrichir ou encore exporter les références bibliographiques collectées lors du processus de recherche. Il est disponible en extension pour le navigateur Firefox. Zotero fonctionne sur le principe des *translators*, ce qui signifie la nécessité de développer de nouveau *translators* pour donner accès aux différents sites qui ne sont pas encore accessibles et exploitables.

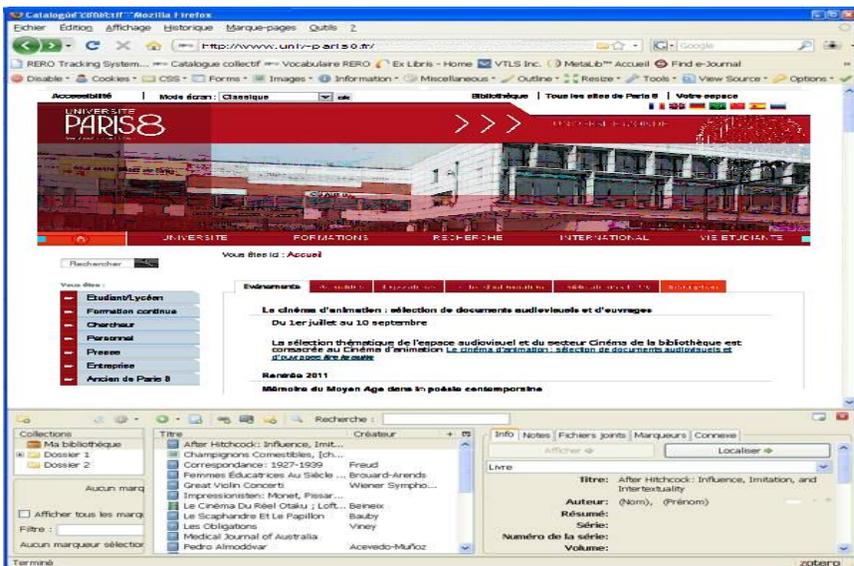


Figure 5. Zotero comme extension au navigateur Firefox (partie inférieur)

Zotero se base sur le développement collaboratif³⁴ pour la mise en place de ces nouveaux *translators* et aussi des styles bibliographiques au format CSL³⁵. L'essentiel de l'outil se résume par sa possibilité à :

³³ George Mason University

³⁴ Par la communauté des utilisateurs

³⁵ Pour « Citation Style Language »

saisir manuellement et ajout de références : possibilité d'importer des références directement à partir de plusieurs interfaces de recherche³⁶ ;

organiser et gérer des références : cela consiste à rechercher, gérer les doublons et les groupes, relier les références et gérer les notes et les fichiers joints ;

créer des bibliographies : cela consiste à la sélection et à l'insertion des références dans le document en cours (Word et OpenOffice) ;

partager et diffuser ses références : par le biais d'un service Web et permettre à d'autres utilisateurs d'y accéder selon les droits en ajout, en modification et en suppression par l'intermédiaire de groupes partagés.

6. Conclusion

La gestion documentaire personnelle est un défi pour chaque chercheur, étudiant, enseignant ou tout autre utilisateur enquêteur d'informations. Obtenir une représentation précise d'une thématique et faire une veille nécessite la mise en place d'un processus de recherche maîtrisé et d'une gestion documentaire complète. Une démarche qui nécessite la connaissance des outils et des services disponibles, et d'en choisir la solution adéquate à adopter pour sa stratégie de gestion.

Les solutions pour la gestion documentaire, pour la gestion des signets et des références, ont fait une véritable percée depuis quelques années, leur intérêt ne se résume plus à des fonctionnalités primaires, mais élargissent leur champ d'action à d'autres, pour répondre à l'évolution des pratiques documentaires, aux exigences des utilisateurs dans ce domaine, à leurs centres d'intérêt.

Cet intérêt s'est largement consolidé avec les innovations que connaissent le Web et l'apparition des services en ligne ou la migration de certains outils vers le réseau, essentiellement pour répondre de nouveaux besoins et faire apparaître à l'occasion de nouvelles pratiques dans le champ de la recherche documentaire... L'offre actuelle permet d'effectuer plus facilement un choix intéressant soit au niveau de la simplicité ou de la richesse des fonctionnalités et surtout dans un cadre

³⁶ Sites et bases compatibles

universitaire pour étudiants et chercheurs. La plupart des logiciels de gestion documentaire, de signets et des références bibliographiques en open source, présentent et suivent les tendances du moment, essentiellement relatives à la collaboration et la syndication des contenus (flux RSS) ou pour les extensions (*plug-in*³⁷ pour navigateur).

Les solutions propriétaires restent toujours un refuge pour le monde professionnel avec des avantages comme la garantie d'un suivi technique et des inconvénients comme le coût des licences... Les logiciels libres permettent à moindre coût et avec peu de manipulations intermédiaires³⁸, de constituer des solutions assez intéressantes et efficaces pour une petite ou une moyenne structure, que ce soit commerciale, éducative ou recherche.

7. Bibliographie

[BEL 99] BELISLE C., ZEILIGER R., CERRATTO T., S'orienter sur le Web en construisant des cartes interactives avec Nestor, Actes H2PTM 99, Hermès, 1999

[BOU 02] BOUHAÏ N., Lire, réécrire et partager le savoir sur le Web : problèmes et solutions, Thèse de Doctorat en Sciences de l'Information et de la Communication, Université Paris VIII, 2002

[BOU 05] BOUHAÏ, N., MORVAN, F., REGOTAZ, D., Un environnement d'assistance à la navigation et à la restructuration d'informations sur le Web. Conférence IHM'05, pp. 331-334, Toulouse, 2005

[BOU 06] BOUHAÏ, N., MORVAN, F., Apprentissage : Une approche orientée espaces de connaissances expertisées, Colloque international E-prospectives et territoires de la connaissance - Les Journées de THOT - 5, 6, 7 octobre 2006 – Albi, France.

[BOU 04] BOUHAÏ, N., MORVAN, F., HNLS : une approche

³⁷ En informatique, un plugin ou plug-in, aussi nommé module d'extension, module externe, greffon, logiciel en France, add-in ou add-on est un logiciel qui complète un logiciel hôte pour lui apporter de nouvelles fonctionnalités (définition Wikipédia)

³⁸ Le format BibTeX est présent dans la majorité des outils, cela permettra l'importation et l'exportation des données inter-logiciels.

constructiviste de connaissances, Actes du colloque TICE2004, Compiègne, 20-22 octobre 2004, pages 219-222.

[CAD 00] CADIZ J. J., GUPTA A., GRUDIN J., Using Web Annotations for Asynchronous collaboration Around Documents - Actes des ACM 2000 Conference on Computer Supported Cooperative Work (CSCW2000), Philadelphia, Pennsylvania, USA, 2000

[DUS 00] DUSSAUX G., PECUCHET J-P., IronWeb : une architecture distribuée pour la création de connaissances pédagogiques sur le Web, in Technologies de l'Information et de la Communication dans les Enseignements d'ingénieurs et dans l'Industrie, Troyes, 2000

[MAR 98] MARSHALL C., Toward an ecology of hypertext annotation – In : Acts de 9th ACM Hypertext and Hypermedia Conference, Pittsburgh, 1998

[REN 09] RENEAR A. H., et PALMER C. L., Strategic reading, ontologies, and the future of scientific publishing. Science (New York, N.Y.) 325, n° 5942, 2009

[SAL 01] SALEH, I., Web Sémantique, Dynamique et Coopératif, Habilitation à diriger des recherches, Université Paris VIII, 2001

[TEN 08] TENOPIR C., DONALD W. K., Electronic Journals and Changes in Scholarly Article Seeking and Reading Patterns. D-Lib Magazine 14, n° 11, 2008

Is Digital Humanities changing the learner-instructor interaction: an inquiry of social media use in archival education in Algeria?

Behdja Boumarafi

Maître de conférences

Institut de bibliothéconomie et documentation

Université de Constantine² Abdelhamid Mehri

Abstract

Digital technology is changing the way we learn, interact, work and entertain, for its unlimited potential in penetrating all spheres of life. The digital revolution is transforming education industry worldwide. In recent years, extensive debate and research are exploring digital humanities(DH), focusing on developing a clear understanding of its capabilities as a platform for making social sciences and humanities applicable to the cyber environment of the twenty first century. The widespread use of social media supported by a rapid growth of the digital culture is making learning ubiquitous by creating, capturing and sharing knowledge. This is enhancing students' engagement and learning efficiency. It is also improving the learner-instructor interaction by engaging students in a more meaningful participation in their own education and academic achievement. To accomplish this, both students and instructors need to have the skills and expertise in capturing the positive effect of the digital technology and engaging in the new learning environment.

This paper reflects on the new learning environment supported by the digital technology.

It discusses how students at the department of archives at the University of Constantine 2 build on the digital capabilities of social media to create and share knowledge.

Keywords: Digital technology, digital humanities, archival education, social media, students' engagement, university of Constantine2.

Introduction

As societies are moving towards the digital era there is a greater need for harnessing the technological advancements of the web and promoting its use in instructional and learning activities. In the last few years, extensive discussion and heated debates are exploring such use. Much of that focused on developing a clear understanding of the capabilities of information and communication technology as a platform for enhancing instruction and learning. Digital humanities are putting academic institutions under pressure for organizational changes built on digital technology to make their education system applicable to the cyber-environment. However, social sciences' programs face innumerable challenges in nurturing and managing the impact of such developments. With the advent of the Internet technology and its gradual penetration in developing countries, education re-engineering is needed to optimize the positive effect of the web technology and its growing applications in instruction and knowledge creation, sharing and transfer. Social media emerged as one of platforms for such endeavors for their unlimited potential of making infinite amount of data available to learners. For instance, facebook has more than 1 billion users all over the world. They create and share around 30 billion pieces of content every month. Therefore, the use of digital technology in form of social networks is growing at a scale that is threatening and at the same time promising. On one hand, they have the potential for global involvement of institutions; on the other hand, the question of how to deal with the big-data generated by myriads of users will soon become inevitable and need to be resolved. This requires new skills and expertise that may not be applicable to social sciences and humanities programs such as archival education. Nonetheless, this is in a state of transition as a result of several factors specifically, economic growth and gradual application of ICT in various activities namely: digital preservation, electronic record management and several other areas to optimize the positive effect of the web technologies. The heavy use of digital technology is changing every sphere of life including education where rigid learning models do not work in the digital era that is promoting flexible systems supported by technological tools. In

retrospect, new communication technologies and research data infrastructure are now appreciated by humanities researchers and learners, enriching the connections within the academy and powering the linkages of content and data (Owen, S. et al, 2014a.). Students in Algerian universities are creating and sharing contents through social media, thus engaging in their own learning in a more meaningful way.

The Institute of Library and Documentation was founded in 1983 it developed gradually from a small institute to the status of national institute with a students' population of about 1200 reading for a BA, Master and Doctorate. There are 45 full time faculty members and a number of part timers serving in institute's two departments namely: library science and archival science. The latter has a total 200 students.

This paper explores the use of social media by students in the department of archives at the University of Constantine 2 (Algeria) to bridge the gap between digital humanities and archival education. It investigates the type of social media they use, the types of activities they are likely to use social media for. It looks at how students are investing their technological skills and learning time, for better academic achievements and examines factors affecting their use.

Archival education in Algerian Universities

Archival science came into the mainstream of higher education in Algeria in 2008. However, it started as a single course long before that, offering archives coursework in the library science bachelor program. In 2008, it grew into a standalone program offering bachelor, master and doctorate degrees to a growing number of students. Courses on archives are designed to educate and train a new generation of archivists qualified to take on different roles and responsibilities for archival activities in public and private organizations. To provide an adequate coverage of curriculum content, that encompass academic courses and technical skills that are vital for educating professional archivists is important in building a stronger archival profession in Algeria as well as in fostering professionalization among the younger generation of graduates, who are required to work in the digital era. Therefore, the curriculum is constructed to serve the archival collections in the country, give a new perspective to the profession and open doors of opportunities for archivists to acquire the required expertise to accommodate the technological requirements of the digital

environment to contribute to essential understandings for the development of future archival systems and technologies that operate at a global level. (McKemmish, Gilliland- Swetland, & Ketelaar, 2005). This depends on the extent to which the program content and characteristics respond to the needs of digital humanities. At this point, it is necessary to developing clear strategies for integrating into the digital arena by:

- Developing and providing access to new and innovative learning and instruction tools
- Tapping into the wide ICT applications to provide access to new formats of education packages and develop a common vision of the demands of the digital humanities.

The rapid development of the internet technology, specifically the digital humanities has emerged as the driving forces that are reshaping the global environment within which social sciences and humanities including archives are operating. Such as IT savvy user demographics, (internet generation) complex user needs, changing collection formats (from paper to digital), increased use of networks, global pressure for sharing knowledge among others. Curriculum development of archival education has been carried out a number of times to include course that have a direct impact upon the development of the field of archival science in relation to digital humanities. Courses like: Electronic archiving, Application of new technology to archives, Archive digitization, Electronic record management, Archival institutions through the web, Metadata & archives, Digital technology & archives that are featured in the program are direct applications of digital humanities to archives.

Literature review

The literature cites four dimensions of learning styles using the social web; these are:

1. Active versus reflective: or trying first then think;
2. Sensing versus intuitive: learning the facts or discovering them;
3. Visual learning versus verbal: using visual material or verbal explanations;
4. Sequential versus global: starting by understanding the linear steps or getting the big picture first. (Mavropalias, n.d.)

Modern education systems are shifting the emphasis from instructors towards learners who take part in their own learning.

Students in the department of archives are using social media, specifically, facebook to collaborate with one another in group projects and term papers, they exchange lecture notes, previous exam questions and other learning material related to their course. One possibility afforded by social media is the ability for students in the Humanities to use and create linked data through open systems and through socially constructed linkages which are driven by the perspectives and understandings of individuals (Owen,S.et al. 2014,b.) The ubiquitous presence of social media is generating big-data outputs and has attracted researchers to study both positive aspects and concerns of using such tools in various settings offering new and various ways of using computers or/and mobile devices. (Paliktzoglou and Suhonen, 2014).

As education institutions are embracing social media there is a need to optimize the positive effect of such technologies to bring them into pedagogy to make instruction and learning active and applicable to the cyber environment of the new millennium. (Boumarafi, 2015a.). For a generation immersed in a world of evolving technologies where internet applications, specifically, the web 2.0 tools are having a considerable impact on creating a technology –driven culture in every society.

Kirshenbaum (2010) states that digital humanities are a social enterprise, it is a network between people who jointly research together, argue, compete and collaborate. Students all over the globe are using social media capabilities to create and share content, exchange ideas and establish networked communities. As a result, a huge amount of digital-born data is being created. That is, in a way promoting the digital humanities movement through innovation in humanities scholarship which still needs more investment while recognizing the challenges of infrastructure and staff (Spiro, 2011) in compiling, organizing and preserving social media content for future use. They need to create new digital tools for data warehousing and text mining. The current approach to digital humanities in social sciences is still limited in terms of scope and research projects. However, as research in this area increases, demand for digital scholarship will become inevitable (Green, 2014) in a new intellectual space aiming for a global impact and internationalization (Zorich, 2008). This has challenged the status-quo of social sciences and humanities as it requires new skills and competencies to engage in digital humanities.

Pahl (2003) studied the evolution and change in web-based teaching and learning environments with the focus on four

perspectives including content, format, infrastructure and pedagogy. The author concluded that lack of standardized technology, its limited life expectancy and cost are among the major problems facing teaching and learning environments that are struggling to keep up with the constant changes in technology and theoretical advances in education.

In 2009, Pannapaker called digital humanities “the next big thing in a long time, because the implications of digital technology affect every field” including arts and social sciences as a result dozens of grants have been awarded to projects in digital humanities, focusing on the application of computing technology to humanistic inquiries and on humanistic reflections on the significance of that technology (Sula, 2013) to develop a clearer understanding of the capabilities of such technology as a new platform in leveraging arts and social sciences efficiently and addressing large scale participation in the creation of digital content and tools.

Ellison, Steinfield & Lampe, (2007) observed that Facebook supports resource sharing by establishing the social foundation between students and their peers. In essence the advent of social networking technology is also the advent of new learning systems and a rapid growth in educational technology. Although, social networks were not initially created for education purpose, Paliktzoglou, Stylianou, and Šohenen (2014) found evidence that Google apps can support pedagogical activities by increasing students’ engagement and team work. Therefore, it is important for archival students to learn how to integrate evolving technology into learning strategies; not just for technology’s sake, but for the added value that these tools already familiar to learners provide. (Brotherton, 2011). Boumarafi (2015b.) investigated the use of social media by Algerian students and the extent to which they use it for academic purpose found that facebook is the most popular.

Methodology

Documented literature was examined to design a questionnaire for this study. A pilot survey was conducted with a small group of archives and library sciences students to assess the weaknesses if any and strength of the questionnaire. Based on their suggestions the instrument was revised and then finalized. Faculty members were approached for permission to distribute the questionnaire during their class session. This allowed greatest accessibility to the target population that consisted of

master students who were in class on the day the questionnaire was distributed. 120 completed questionnaires were collected, out of which 73 (60.8%) were used in the study the rest were considered unusable. The majority of these respondents were female (n=43; 31.39 %) and male represent (n= 30; 21.9%), in fact the number of female students in the institute in general outnumbers the male. All respondents have a laptop, a smart phone and internet connection at home. This means that they have the possibility to respond to the digital humanities material requirements.

Fidings

Preferred social media

Respondents were asked which social network they prefer to use. Their preferences are summarized in table 1. As expected, all respondents 73 (100%) gave the top rank to Facebook at a mean of 4.10. Previous studies also identified Facebook as the most frequently used social network site (SNS). YouTube is ranked second in preference with a mean of 4.00. Respondents put Twitter in third position in terms of importance with a mean of 3.85, followed by LinkedIn which scored 3.64. Skype that respondents use to keep in touch with family and friends at home and abroad was put in fifth position with a mean of 3.63. MySpace is ranked last with a mean of 3.10

Type of social media	Mean	Sdt. Dev.	Rank
facebook	4.10	1.163	1
YouTube	4.00	1.005	2
Twitter	3.85	1.078	3
LinkedIn	3.64	1.106	4
Skype	3.63	1.123	5
Google+	3.10	1.265	6

Table 1. Preferred social network sites (n=73)

Activities carried out using Social media

Respondents were asked to specify academic activity they carry out using social media and indicate the importance of each. The respondents' responses are presented in the form of means and standard deviation in table 2.

activities	Mean	Std. Dev.	Rank
Discuss group projects	4.50	.996	1
Share assignments and course work	4.31	.917	2
Share files and lecture notes	4.23	.927	3
Create content	4.17	1.116	4
Exchange ideas	4.12	1.029	5
Join academic discussion forums	4.06	1.165	6
Make a presence in the cyber-space	4.03	1.006	7
Improve foreign language skills	3.99	1.109	8
Self regulated learning	3.94	1.106	9
Improve communication skills with students abroad	3.81	1.092	10
Share ideas and promote creativity	3.74	1.094	11
Make contact with faculty members easier	3.50	1.124	12
Enhance academic achievement	3.35	1.148	13
Get assignments from faculty members	2.99	1.220	14
Send assignments to faculty members	2.80	1.265	15
Get grades for assignments completed	2.36	1.100	16

Table 2. Activities carried out using social media (n=73)

From the results exhibited in table 2 above, the sample groups of the study consider seven activities as important at various levels. However, "discussion of group projects" is considered as the most important and gets top rank with a mean of 4.50. They also placed high importance on sharing assignments and course work (mean=4.31), followed by "share files" and lecture

notes”(mean=4.23), “create content”(mean=4.17), “exchange ideas”(mean=4.12), “join academic discussion forums”(mean=4.06), and “make a presence in the cyberspace”(mean=4.03). All these activities are very close to very important. Learning a foreign language especially English became a very important skill for the 21st century learners. Therefore it seems logical that respondents considered “improve foreign language skills” as the next important activity they carry out (mean=3.99). “Self regulated learning” was also perceived as important (mean=3.94). In the same manner, the study showed that the perceived importance of “improve communication skills with students abroad” was evaluated with a mean of 3.81. “Share ideas and promote creativity”(mean=3.74) this supports the claim that students use social media to create content. Respondents agreed that social media make contact with faculty members easier (mean=3.50). Enhance academic achievement”(mean=3.35) indicate that students benefit from the use of social media technologies that have positive effect on their academic performance and growth. The last three activities namely: “Get assignments from faculty members”, “send assignments to faculty members“, and” get grades for assignments completed” were given less importance with means less than 3; 2.99, 2.80, 2.36 respectively. These results reveal digital humanities acceptance and adoption by Archival science students expressed in a good use of social media in a numbers of academic activities.

Discussion

Interest in Digital Humanities is growing because of the belief that digital technology is becoming essential for socialization, work and study. Arguably, it seems certain that learning and research in humanities and social sciences will be affected by their evolution into digital forms. Therefore, there is a need to understand what motivates archival students at the institute of library and documentation science to use social media as digital tools to participate in their own learning and enhancing the instructor-learner interaction in the process. The study found evidence that facebook is the most used tool and received top score. This result is supported by previous studies. MySpace is the least popular among participants. In relation to students activities carried out using social media, the study revealed that the top activities are of academic type between peers. This helps in the deployment of DH in archival studies. Such deployment impacts archivists in the way they study and represent

knowledge with direct improvement in instruction and learning efficiency. Surprisingly, interaction with faculty members is unexpectedly low. No doubt, using DH along with social networks increase understanding of digital techniques in humanities and social sciences and yield a better collaboration between learners and instructors in sharing knowledge and ideas and learn from each other's opinions. As digital tools, that is what social media is about. Whether creating, collecting, interpreting and sharing knowledge through social media tools.

Conclusion

The study investigates the use of social media as digital tools for enhancing the instructor-learner interaction in the department of archives at the institute of library and documentation science. Archival students use such tools mainly facebook for different purposes to create new spheres of interactive learning with ultimate connections to global networks to share knowledge with peers abroad. The use of social media enabled them create content and share it with peers. Surprisingly, the results of this study show that respondents give little importance to the instructor-learner connection. Overall, our study revealed that, although there is evidence of the use of digital humanities by archival education for independent learning. The study reported changes in curriculum development to engage in DH. Yet little empirical work has been done to envision exactly how to construct a framework to facilitate collaborative actions between technology experts and archivists to incorporate DH at the pedagogical level as tools for archival teaching and interpretation. Only then the curriculum will reflect the digital knowledge and tools required.

Références

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" social

capital and college students' use of online social network sites. *Journal of Computer Mediated*

Communication, 12(4), 1-26.

Boumarafi, B. (2010). Strategies for the delivery of e-information services to support the e-learning environment at

the University of Sharjah. *The Electronic Library*, 28(2), 276-285.

Boumarafi, B. (Summer 2015). Social media use in Algerian Universities: University of Constantine 2

case study. *The IAFOR Journal of Education: Technologies & Education Special Edition*
31-41

Brotherton, P. (2011). Social network enhance employee learning. *ABI/INFORM Complete*, 65(4),18

Mavropalias, K. (nd). Social bits: personality and learning style profiling via the social web

Retrieved 10/03/2016 from: <http://www.Iconof.com/blog>

Green, .(2014). *Digital humanities: a new model of scholarship in a new intellectual space*. In supporting digital humanities for knowledge acquisition in modern libraries. Retrieved 10/03/2016 from: [http://books/google/books? isbn=1466684453](http://books/google/books?isbn=1466684453). p.219.

McKemmish, S., Gilliland-Swetland, A. & Ketelaar,E. (2005). Communities of memory: Pluralising. *Archival Research and Education Agendas Archives and Manuscripts* 33, 146-174

Owen, S.; Verhoeven, D.; Horn, A.; Robertson, S. (2014). Collaboration success in the dataverse libraries as digital humanities research partners. In: proceedings of IATUL conferences. Paper1. Retrieved 20/03/2016 from: <http://docs.lib.purdue.edu/iatul/2014/openaccess/1>

.Pahl, C. (2003). Managing evolution and change in web-based teaching and learning environments. *Computers and Education*. 40: 99-114.

Palktzoglou, V., Stylianou, T., & Suhonen, J. (2014). Google educational apps as a collaborative learning tool among computer science learners. In: *Assessing the Role of Mobile Technologies and Distance Learning in Higher Education* Edited

by: P. Ordóñez de Pablos, R.D. Tannyson, M.D. Lytras. 272-296
IGI Global isbn: 9781466673168.

Pannacker, W. (2009, December 28). The MLA and the digital humanities. *Chronicle*

of Higher Education. Retrieved 13/03/2016 from
<http://chronicle.com/blogPost/The-MLAthe-Digital/19468/>

Spiro, L. (2011). This is why we fight: defining values of digital humanities. Retrieved 10/03/2016 from:
<http://www.nyu.edu/projects/Senger/CDH/spiro.pdf>

Sula, C.A. (2013). Digital humanities and libraries: a conceptual model.

Journal of Library Administration vol. 53, pp. 10-26

Doi:10.1080/01930826.2013.756680

Kirschenbaum, M.G. (2010). What is digital humanities and what's it doing in English department. Retrieved 13/03/2016 from:
<http://mkirschenbaum.files.wordpress.com/2011/03/ad-final.pdf>

Zorich, D.M.(2008). A survey of digital humanities centers in the United States. Retrieved 13/03/2016 from:
<http://www.clir.org/pubs/reports/pub143.pdf>

La médiation numérique et l'avenir des métiers de l'information-documentation face à l'inflation du document numérique

Mor DIEYE

Enseignant-chercheur en Sciences de l'information
documentaire

École de Bibliothécaires, Archivistes et Documentalistes –
UCAD

Mots-clés :

Archives – Document – Information – Médiation – Numérique –
Usager – Web.

Résumé

L'objectif de cette communication est de tenter de montrer comment les métiers de l'information-documentation pourraient adopter les fonctions de la médiation numérique dans les pratiques professionnelles ? Cette question est d'autant plus importante dans un contexte où d'aucuns se posent la question d'une éventuelle supplantation de ces métiers par les technologies du numérique. A ces questions s'ajoutent celle de l'inflation informationnelle due à une tendance de plus en plus accrue au « tout numérique ». Il s'agira ainsi de voir si l'intégration de la médiation numérique par les professionnels pourrait aider à une meilleure optimisation de la diffusion de l'information numérique ?

Introduction

Face à l'inflation de l'information numérique en raison de l'explosion des réseaux du web et des médias sociaux, à l'évolution des usages et la transformation du rapport entre les

utilisateurs et les moyens d'accès aux contenus informationnels, il est plus que jamais nécessaire de s'interroger sur les types d'information à mettre à la disposition des usagers. Le flux d'informations de toutes sortes qui circulent sur la toile virtuelle doit davantage inciter les professionnels de l'information-documentation à davantage de rigueur dans le tri des contenus à proposer aux usagers. Cela est d'autant plus nécessaire que sur certains sites web professionnels spécialisés, soit en information documentaire, soit en information d'actualité, réputés sérieux et crédibles, il est maintenant difficile de faire la différence entre ce qui relève de l'information scientifique fiable et sérieuse, et ce qui relève de l'information brute, non validée voire tronquée. Le paysage de l'information numérique change non seulement le comportement des usagers, mais réoriente aussi le rôle des services traditionnels de recherche en bibliothèque ou en service d'archives. Les usagers des unités documentaires sont désormais devenus des consommateurs d'information qui passent instantanément de moteurs de recherche commerciaux à des sites de réseaux ou médias sociaux, des wikis, des signets et des services électroniques fournis par ces mêmes unités documentaires pour leurs besoins d'information.

Avec une tendance au « tout numérique » que professent certaines personnes dont, parfois, des spécialistes des technologies de l'information et de la communication, il est légitime de se demander si les métiers de l'information-documentation échapperont à cette nouvelle donne de l'essor numérique. Aujourd'hui, aucun corps de métier ne se soustrait de ce qu'il est convenu d'appeler l'« inflation numérique » de la sphère informationnelle. On ne peut nier l'importance et l'utilité du numérique dans les pratiques professionnelles ou personnelles quotidiennes, et il faut convenir que la dématérialisation ou la conservation et la gestion électronique devient de plus en plus un phénomène de société dans tous les secteurs d'activités. Cela doit inciter à réfléchir aux conséquences qu'une telle conception du numérique peut avoir sur certains métiers dont ceux de l'information-documentation.

Ainsi, notre communication sera axée sur deux points essentiels. Dans un premier, nous tenterons d'aborder la nouvelle approche du traitement du patrimoine documentaire face à l'inflation du document numérique et en second lieu, nous essayerons de voir comment peut-on envisager l'avenir des métiers de l'information-documentation avec l'intégration de la médiation numérique dans les pratiques professionnelles ?

Nouvelle approche du traitement du patrimoine documentaire face à l'inflation du document numérique

Il faut tout d'abord préciser que le concept de « patrimoine documentaire » recouvre aussi bien l'information contenue sur support papier que celle conservée sur support numérique. Mais, il sera plus question ici du patrimoine documentaire numérique et ses moyens et outils de traitement pour un accès facile et optimisé. L'utilisation croissante et rapide des contenus numériques dans les activités administratives, de recherche et dans la diffusion des connaissances est une caractéristique principale que les métiers de l'information-documentation prennent désormais en considération. Le rôle traditionnel de ces métiers qui consistait à aider à naviguer dans de grands systèmes bibliographiques complexes des bibliothèques ou dans les instruments de recherche des mines d'informations que conservent les fonds des services d'archives semble menacé par des services des médias du web qui offrent un choix informationnel quasiment illimité et qui contournent les services traditionnels de gestion du patrimoine documentaire.

Dans le cadre de la médiation numérique, les contenus informationnels à mettre à la disposition des publics doivent désormais passer par l'étape de ce que l'on pourrait appeler du « désherbage numérique » à l'instar du désherbage classique traditionnel pratiqué dans certaines unités documentaires comme les bibliothèques. Le désherbage physique « *consiste à retirer des rayonnages en magasin ou en libre-accès les documents qui ne peuvent plus être proposés au public. On parle également d'élimination, d'élagage, de retrait, de révision ou de réévaluation* »³⁹. En d'autres termes, parmi les masses d'informations contenues dans les collections numériques, il est indispensable d'opérer des mises à jour et des révisions critiques qui permettent d'évaluer leur intérêt administratif, scientifique et historique qui justifieront leur conservation ou leur retrait des fonds numériques.

Les bases de données des établissements documentaires accessibles soit, par le biais d'un portail virtuel local (à partir d'un serveur), soit au travers d'un réseau Internet, sont souvent surchargées d'informations désuètes dont l'actualisation ou tout

³⁹ DAVID Stéphanie. *Le désherbage physique*. Fiche créée le 17 janvier 2008. [En ligne] sur : <http://www.enssib.fr/bibliotheque-numerique/document-1735>. [Consulté le 20-03-2016].

simplement le retrait devrait intervenir depuis très longtemps. En guise d'exemple, nous considérons une exposition virtuelle réalisée par un service d'archives sur un thème relatif à une commémoration occasionnelle, qui peut encore être visible sur le site web de ce service plusieurs mois, voire plusieurs années après la fin de l'exposition. Pendant toute cette période, les contenus de cette exposition occupent inutilement un espace qui pourrait être consacré à d'autres informations beaucoup plus intéressantes pour les usagers à condition de sauvegarder sur d'autres espaces car l'historique des expositions peut être un choix intéressant. Le désherbage numérique permet non seulement d'offrir au public des documents attrayants, plus actuels, mais optimise aussi les espaces virtuels de conservation et de gestion. De plus, l'essor des moyens de communication et de d'échange interactif par message instantané fait que les sites web qui hébergent les documents numériques intègrent de plus en plus les médias sociaux (Forum, Blog, Wiki, Agrégateur, FAQ collaborative, Service de géolocalisation, etc.) qui sont à la fois des espaces d'instruction pédagogique, de détente et de convivialité. La mise à jour des fonds numériques permet aussi de mettre en valeur les nouvelles collections et de susciter ainsi les demandes du public.

Au niveau des services d'archives, le terme de désherbage correspond au « tri » ou « triage » pratiqué dans le traitement des fonds d'archives. La nuance qu'il serait utile de relever entre « désherbage » classique en bibliothèques et « triage » physique en archives est qu'en effet l'opération de désherbage intervient bien en aval après le traitement documentaire des ouvrages (et surtout quand des problèmes d'espace de conservation se posent) alors que le second concept concerne le traitement en amont des documents d'archives. En effet, le travail de tri s'opère lors du classement des archives, c'est-à-dire, avant leur conservation et leur communication aux usagers.

En ce qui concerne l'inflation numérique dans le domaine des archives, nous parlerons donc du « tri numérique » des documents pour y faire face. Cela consiste – s'il s'agit des documents qui ne sont pas originellement numériques, c'est-à-dire qui sont numérisés –, à vérifier les délais légaux d'utilité administrative ou historique prévus dans la charte de conservation des archives. La connaissance de ces délais lors des opérations de numérisation des archives devrait normalement éviter de dématérialiser des types de documents dont la date d'élimination est dépassée ou est très proche. S'il

s'agit des documents d'origine électronique, leur mise à jour dépendra aussi bien des délais légaux de leur durée de conservation, que de l'intérêt administratif, légal et historique qu'ils présentent pour le public. Une rare consultation de ces documents sur le web, par exemple, peut être un indicateur pertinent pour décider de leur retrait des fonds numériques mis en ligne.

Tout le flux d'informations numériques disponibles sur le web n'a pas forcément une vocation informative d'un point de vue scientifique ou historique. L'intérêt d'un contenu informationnel se mesure à travers ses éléments présentant des aspects à la fois d'apprentissage et de transmission. Le cycle de transmission de l'information est aujourd'hui accentué et accéléré par l'outil numérique. Au-delà de la transmission numérique, la dimension éducative de l'information ne doit pas non plus être négligée, car *« les documents numériques possédant une dimension d'apprentissage sont majoritairement non institutionnels. Le paradigme informationnel évacue fréquemment le document ou néglige son rôle clef dans la transmission de savoirs et de savoir-faire en privilégiant une logique de flux »*⁴⁰. Une information numérique doit cependant être considérée davantage au regard de l'ensemble de ses dispositifs informationnels qui prennent en compte les demandes des usagers. L'importance du contenu informationnel à servir aux usagers démontre l'utilité et la nécessité d'insister sur la qualité du support informatif, c'est-à-dire, du document numérique.

Compte tenu de la libre circulation informationnelle à laquelle nous assistons aujourd'hui, il est aussi primordial de penser à la protection des informations sensibles ou tout simplement personnelles. En effet, par imprudence professionnelle, des documents numériques confidentiels pourraient facilement se retrouver sur les réseaux ou médias sociaux et être accessibles au grand public. A cela s'ajoute un système de contrôle et de fichage des informations que certains utilisent pour lutter contre la cybercriminalité ou le terrorisme. Un exemple en date du mois de juin 2013 concerne les confessions ahurissantes d'Edward Snowden, un ancien employé de 29 ans de la NSA (National Security Agency) sur les méthodes peu orthodoxes utilisées par cette agence. La NSA est un organisme

⁴⁰ LE DEUFF Olivier. Logiques documentaires et enjeux éducatifs. [En ligne] sur : <http://www.cairn.info/revue-document-numerique-2012-3.htm>. [Consulté le 20-03-2016].

gouvernemental des États-Unis, responsable du renseignement d'origine électromagnétique et de la sécurisation des systèmes de communications et de traitement des données du gouvernement américain. Mais, les abus supposés ou réels que commettrait cette agence en matière de surveillance sont dus aux pouvoirs disproportionnés qu'elle s'est octroyée. Car, selon Jérémie Zimmermann⁴¹ « *le paragraphe 1881a de la loi amendant l'US Foreign Intelligence Surveillance Act (FISA), voté par le Congrès en 2008, permet aux services d'intercepter toutes les communications de citoyens non américains, pour quelque motif que ce soit, sans mandat judiciaire et donc sans laisser de trace* »⁴². Cette légitimité législative accordée par le Congrès américain à l'agence lui permet de collecter des informations personnelles et confidentielles sans aucun contrôle. En effet, selon toujours M. Zimmermann « *une présentation interne de la NSA nous apporte finalement la preuve irréfutable que toutes les communications et les données personnelles des utilisateurs de Microsoft, Yahoo !, Google, Facebook, AOL, Skype, YouTube, ou encore Apple ont été massivement récupérées par la NSA et le FBI ces dernières années, sans aucune forme de contrôle démocratique* »⁴³. Cette situation entraîne une certaine fébrilité dans la protection des données sensibles et stratégiques de certains gouvernements des pays occidentaux. C'est le cas de la Russie qui comptait retourner aux méthodes anciennes de protection des informations de ses services spéciaux. En effet, selon l'AFP du 11/7/2013 « *après les révélations de l'informaticien américain Edward Snowden, l'un des services spéciaux russes a décidé d'acheter des machines à écrire pour taper des documents secrets, écrit le quotidien russe Izvestia. Le FSO, Service fédéral de protection, issu de l'ancien KGB, a lancé un appel d'offre pour l'achat de 20 machines à écrire* »⁴⁴.

Une bonne politique de gestion du document numérique doit non seulement assurer la bonne conservation et la sécurité des

⁴¹ Cofondateur et porte-parole de la Quadrature du Net, une organisation de défense des droits et des libertés des citoyens sur Internet.

⁴² GRALLET Guillaume. *Le Point.fr* du 10/06/2013. [En ligne] sur : http://www.lepoint.fr/technologie/zimmermann-faire-confiance-a-google-ou-facebook-c-est-etre-a-poil-sur-internet-10-06-2013-1678994_58.php. [Consulté le 21-03-2016].

⁴³ *Ibid.*, GRALLET Guillaume. *Le Point.fr* du 10/06/2013.

⁴⁴ AFP. *Dépêche* du 11/7/2013. [En ligne] sur : <http://www.france24.com/fr/20130711-service-special-russe-revient-a-machine-a-ecrire>. [Consulté le 21-03-2016].

données, mais aussi d'une bonne analyse de la valeur que celles-ci présentent pour le public. Il s'agira de réfléchir aux parades permettant de contourner les menaces qui pèsent sur les systèmes d'information. Pour cela, il faudra procéder à une veille des principaux outils de prévention à mettre en œuvre. Mais, au-delà des spécificités qui caractérisent un établissement de gestion du patrimoine, les réseaux et systèmes d'information sont devenus, de manière générale, les centres nerveux de toutes les institutions quelles qu'elles soient.

Étant donné que les sources numériques constituent le cœur d'un système de médiation numérique, il est essentiel d'en déterminer, lors du traitement, la valeur informative qui nécessite leur protection contre des attaques malveillantes. Avant toute mise en ligne, il est fondamental d'analyser l'intérêt informationnel des données, leur criticité et leur sensibilité par rapport à l'utilisation et l'exploitation qu'en feront les usagers. En effet, quel que soit le niveau de performance des moyens mis œuvre, la technique ne pourra pas tout devant les nouvelles menaces et l'inflation documentaire. Il faut donc rationaliser la production documentaire numérique en procédant par un tri qualitatif qui identifie les documents vitaux répondant aux besoins des usagers. Il est aussi important de se réapproprier les bonnes pratiques professionnelles optimales qui garantissent à la fois une bonne protection sécuritaire et une meilleure optimisation de la production.

Pour illustrer la vulnérabilité de certains systèmes d'information, nous pouvons considérer l'exemple des bases de données qui contiennent les données les plus sensibles d'une entreprise. Ces données peuvent concerner des dossiers financiers, des informations relatives à des cartes de crédit, à des citoyens, à des clients, etc. Et, si nous nous référons à l'actualité sur la cybercriminalité, nous remarquerons que ces données citées constituent la première cible des violations de sécurité. En conséquence, les politiques de sécurité des données des entreprises sont maintenant pour l'essentiel concentrées sur les menaces, les attaques externes et sur la rationalisation des processus de conformité en matière de protection des actifs informationnels les plus stratégiques.

Pour donner un ordre d'idée sur cette fraude de données, les chiffres de l'IBM InfoSphere Guardium indiquent que *« 95% des cas de violations de sécurité ont lieu sur des bases de données. Le coût moyen de l'atteinte à la protection des données s'élève à 7 200 000 dollars. 88% des entreprises interrogées ont*

déclaré avoir subi au moins une atteinte à la protection des données »⁴⁵.

Il faut enfin souligner que le traitement des données doit prendre en compte les risques qui entourent l'inflation du document numérique par rapport à la fiabilité et à la crédibilité des contenus informationnels qui, en effet, peuvent entraîner une insécurité à la confidentialité des données. Malheureusement, les professionnels de la sécurité des données pensent que les atteintes à la protection des données, des modifications non autorisées des bases de données et des plateformes de partage de fichiers seront encore plus importantes. Ainsi, les mesures de protection doivent davantage être de mise.

1. Médiation numérique et avenir des métiers de l'information-documentation

Quel avenir professionnel la médiation numérique réserve-t-elle à l'archiviste, au bibliothécaire et au documentaliste de demain ? Cette interrogation n'est pas anodine et concerne les transformations récentes, en cours et à venir de ces métiers. Dans les sphères des établissements de gestion du patrimoine, cette question se pose à divers degrés selon les types d'établissements : ici, elle prendra une importance grandissante, là elle ne préoccupera pas encore trop les professionnels car les performances actuelles des outils du numérique ne sont pas encore en mesure de supplanter tout le travail de traitement documentaire de l'information effectué par ces professionnels.

Face au développement de la médiation numérique, il est quand même indispensable de repenser les profils des métiers de l'information-documentation en termes de compétences professionnelles. Cette réflexion doit se faire en dehors de toutes considérations de profils de postes liés à des spécificités conjoncturelles. Elle doit plutôt consister à « *définir un ensemble précis de compétences (savoirs et savoir-faire) et un certain panel de tâches requérant ces compétences, à un certain*

⁴⁵ IBM InfoSphere Guardium. [En ligne] sur : <http://www-01.ibm.com/software/fr/info/rte/bdig/dsp-1-post.html>. [Consulté le 21-03-2016].

niveau de responsabilités et d'architecture relationnelles au sein d'un établissement »⁴⁶.

Les professionnels de l'information-documentation ne doivent pas avoir des appréhensions vis-à-vis du numérique, mais ils doivent le considérer comme un apport technique et intellectuel à leurs métiers. Le management (traitement, gestion, conservation, diffusion, etc.) de l'information documentaire vit, de nos jours, une véritable révolution procédurale. Nous assistons aussi « à l'émergence d'une société de l'information basée sur le développement exponentiel des usages d'Internet et des réseaux électroniques et sur la généralisation du document numérique »⁴⁷. Des référentiels classiques et traditionnels jusque-là utilisés, il faudra passer aux nouveaux référentiels de métiers apportés par la révolution du numérique. En outre, il est aussi nécessaire de changer le niveau des préoccupations professionnelles en adoptant la veille numérique. Enfin, il est également important, comme le dit Jean Michel, d'« adopter de nouveaux modèles de management de l'information, raisonner désormais dans une perspective infiniment plus ouverte et selon des schémas d'interactivité nettement plus affirmés »⁴⁸.

Il ne s'agira donc pas de remplacer un métier par une nouvelle pratique professionnelle (c'est-à-dire, les métiers de l'information-documentation par la médiation numérique), mais plutôt d'adopter de nouveaux modèles et outils de travail et de les positionner dans les organisations professionnelles afin d'en tirer le meilleur parti. Autrement dit, mettre les potentialités du numérique, du multimédia et des réseaux au service de l'information-documentation sans pour autant renier ses missions documentaires originelles.

Tous les métiers connaissent des évolutions provoquées, soit par le développement industriel, soit par des mutations technologiques. Si nous remontons dans le temps, nous

⁴⁶ MERCIER Sylvère [dit Silvae]. *Bibliobsession. La fonction ressource et les métiers de l'information-documentation*. En ligne] sur : <http://www.bibliobsession.net/2010/06/02/la-fonction-ressource-et-les-metiers-de-linformation-documentation/>. [Consulté le 21-03-2016].

⁴⁷ MICHEL Jean. *Les professionnels de l'information-documentation à l'heure du document numérique et des réseaux électroniques*. [En ligne] sur : <http://michel.jean.free.fr/publi/JM293.html>. [Consulté le 21-03-2016].

⁴⁸ *Ibid.*, MICHEL Jean.

constatons que les professions en information-documentation ont subi depuis quelques décennies de profondes transformations engendrées, pour l'essentiel, par le développement des techniques de l'information et de la communication. Jadis, la gestion, la conservation et la communication des documents d'archives s'organisaient essentiellement, pour un service d'archives, autour de l'élaboration (à l'issue du traitement des fonds) des instruments de recherche imprimés. Pour les bibliothèques, la gestion des collections se faisait traditionnellement autour du traitement à la fois du livre et de l'information qu'il contient ; c'est en effet le domaine de la description bibliographique et du catalogage qui fournissent l'ensemble des renseignements matériels et informationnels relatifs à l'accès aux ouvrages. Au cours de cette période – qu'on pourrait qualifier d'« intermédiaire » (parce que se situant entre les débuts anciens des métiers du document et leur essor fulgurant actuel) –, les pratiques professionnelles n'en étaient pas pour autant profondément bouleversées. C'est en effet l'époque qui voit apparaître les grandes bases ou banques de données bibliographiques comme les bases PASCAL, COMPENDEX ou INSPEC, etc. Avec ces bases de données, l'utilisateur disposait déjà, grâce bien sûr, à des logiciels de recherche documentaire, d'un outil qui lui permettait de retrouver facilement des documents. Ensuite, nous assistons à l'avènement du document numérique qui, quant à lui, va beaucoup transformer les pratiques professionnelles de l'information-documentation. Ces bouleversements professionnels emmènent certains à s'interroger sur leurs conséquences futures sur les profils et les principes fondamentaux de la fonction.

L'intrusion du numérique dans la gestion de l'information documentaire ne peut, en aucun cas, remettre en cause le continuum des pratiques professionnelles déjà existantes. Néanmoins, cette nouvelle donne redéfinit les champs d'intervention et les niveaux de responsabilité professionnelle, mais doit aussi permettre de mettre sur pied un nouveau dispositif de certification des compétences des professionnels de l'information-documentation. Le référentiel des professionnels de l'information-documentation appelé « Euroréférentiel » mis au point par l'Association des professionnels de l'information et de la documentation (ADBS) sous la direction de Jean Meyriat, donne des orientations nécessaires à la certification des compétences. Par exemple, dans le groupe du domaine de compétence information, nous trouvons « à la fois les

compétences nécessaires au traitement de l'information et celles qui s'appliquent aux supports de l'information, parce que, dans la réalité du travail des professionnels, ces deux pôles traditionnellement distincts commencent à se rapprocher l'un de l'autre, et aussi pour éviter de donner des arguments à ceux qui s'obstinent à distinguer l'une de l'autre ces deux facettes du travail documentaire »⁴⁹.

La question cruciale qu'il faudra cependant se poser est de savoir si l'arrivée du numérique dans le domaine de l'information-documentation pourrait participer au dépérissement des frontières professionnelles entre l'archiviste, le bibliothécaire et le documentaliste ? En effet, entre ces trois fonctions, le métier de documentaliste semble plus proche des exigences et des caractéristiques du numérique. Cela s'explique par le fait que la matière que traite la documentation est souvent d'origine électronique ou est déjà passée par la phase de la numérisation. Les technologies de la dématérialisation des fonds d'archives et des collections des bibliothèques peuvent favoriser le rapprochement entre les fonctions d'archiviste, bibliothécaire et le documentaliste, mais ne peuvent pas, par contre, supprimer les particularités fonctionnelles et les dimensions systémiques qui les différencient. La tendance est aujourd'hui, par le biais d'un jeu sémantique, de modifier les appellations des fonctions ou d'élargir les profils de l'information-documentation, car « *on évoque de plus en plus des métiers tels qu'experts en veille technologique, spécialistes en bibliométrie ou scientométrie, concepteurs de systèmes d'information ou de produits multimédias ou hypermédias, gestionnaires de documentation structurée, rédacteurs techniques, spécialistes de l'archivage électronique, voir encore « chercheurs » ou « infomètres »* »⁵⁰. On pourrait y ajouter celui d'architecte de l'information dont la mission est d'intervenir « *en amont d'un projet de conception de site web, d'intranet ou de logiciel de grande envergure. Il définit la typologie et la classification des contenus, optimise la navigation et la recherche d'information*

⁴⁹ MEYRIAT Jean. Euroréférentiel-compétences et aptitudes des professionnels de l'I-D. *Association des professionnels de l'information et de la documentation (ADBS)*. Paris : 2004. [En ligne] sur : <http://certidoc.net/fr1/tout.php>. [Consulté le 22-03-2016].

⁵⁰ *Ibid.*, MICHEL Jean. *Les professionnels de l'information-documentation à l'heure du document numérique et des réseaux électroniques*.

en fonction des objectifs et du public visé »⁵¹. Ceci peut être positif si l'objectif visé est simplement d'élargir et d'enrichir le domaine d'intervention du métier, mais sans pour autant vouloir ériger ces spécialités en véritables fonctions exhaustives et définitives de l'information-documentation.

La question de l'avenir des métiers de l'information-documentation en cette ère de l'info-document numérique ne doit pas se poser en termes d'un éventuel remplacement d'une fonction par une autre plus récente ou plus moderne, mais plutôt en termes de finalités, de produits et de services à mettre à la disposition des usagers. Tout cela entre dans le cadre du management de l'information en général. Il s'agira en fait de raisonner autrement par rapport à l'évolution d'un métier et à la nouvelle demande de l'utilisateur que l'essor du numérique lui a, dans une certaine mesure, imposé et aux besoins de ce dernier dans ses pratiques consuméristes de l'information. Toute évolution professionnelle entraîne un coût économique, il faudra donc réfléchir avant tout, aux conséquences économiques des nouvelles propositions techniques et professionnelles que cette évolution apporte à la profession. Se demander si ces conséquences sont positives ou négatives dans la valorisation professionnelle du métier par exemple ? En outre, l'intégration des fonctions du numérique dans le domaine de l'information-documentation doit inciter les professionnels à se mettre davantage dans une logique de l'usage et de la valorisation de l'information que plutôt dans une simple logique de production et de diffusion auprès des publics.

Ainsi, la priorité doit être axée sur l'importance de la valeur ajoutée des produits et services d'information. Cela veut dire que le management de l'information ne peut pas se limiter seulement à l'utilisation d'un outil technologique sophistiqué qui permet un meilleur traitement et un large accès et de diffusion, mais il doit prendre en compte des compétences professionnelles humaines qui vont de l'exploitation scientifique de l'information en passant par des recoupements jusqu'à la conception des produits adaptés à d'éventuelles cibles d'utilisateurs. Cette quête professionnelle de valeur ajoutée s'exerce aussi à travers l'expertise, le conseil, la veille, l'alerte, etc. pour la conception des contenus de l'information.

⁵¹ Le Portail des Métiers de l'Internet. *Architecte de l'information*. [En ligne] sur : <http://metiers.internet.gouv.fr/metier/architecte-de-l%E2%80%99information>. [Consulté le 22-03-2016].

Conclusion

Depuis l'avènement du document numérique dans l'espace de l'information documentaire, des mutations ont été constatées aussi bien dans les pratiques professionnelles que dans les comportements de consommation de l'information des usagers. Cette nouvelle donne doit en effet être prise en compte dans le traitement documentaire des contenus informationnels surtout dans un contexte « inflationnel » de l'information numérique. Cela permettra, non seulement, de mieux répondre aux besoins des usagers, mais aussi de leur offrir un accès plus facile et plus optimale à l'information.

Cependant, l'avenir réside dans un regard lucide de l'horizon qui positionnera le professionnel de l'information-documentation au centre du développement de la technologie numérique. Cette position doit montrer la capacité de ce dernier à produire de la matière documentaire pertinente, à évaluer les contenus des bases de données ou des réseaux du web, à conseiller ses usagers et les orienter vers ce qui a de l'intérêt ou de la valeur informationnelle pour eux. Cela s'appelle en effet donner un sens à l'information tout court. Dès lors, le professionnel se présente comme un partenaire à part entière de l'évolution du numérique avec tout ce que cela comporte de conception, du traitement et de diffusion des connaissances.

Enfin, il est maintenant indispensable qu'une complémentarité et une complicité entre la médiation numérique et à la médiation traditionnelle ou classique se fasse au seul service du patrimoine à préserver et de la satisfaction des usagers.

Bibliographie

AFP. *Dépêche du 11/7/2013*. [En ligne] sur : <http://www.france24.com/fr/20130711-service-special-russe-revient-a-machine-a-ecrire>. [Consulté le 21-03-2016].

DAVID Stéphanie. *Le désherbage physique*. Fiche créée le 17 janvier 2008. [En ligne] sur : <http://www.enssib.fr/bibliotheque-numerique/document-1735>. [Consulté le 20-03-2016].

GRALLET Guillaume. *Le Point.fr du 10/06/2013*. [En ligne] sur : http://www.lepoint.fr/technologie/zimmermann-faire-confiance-a-google-ou-facebook-c-est-etre-a-poil-sur-internet-10-06-2013-1678994_58.php. [Consulté le 21-03-2016].

IBM InfoSphere Guardium. [En ligne] sur : <http://www-01.ibm.com/software/fr/info/rte/bdig/dsp-1-post.html>. [Consulté le 21-03-2016].

LE DEUFF Olivier. Logiques documentaires et enjeux éducatifs. [En ligne] sur : <http://www.cairn.info/revue-document-numerique-2012-3.htm>. [Consulté le 20-03-2016].

Le Portail des Métiers de l'Internet. *Architecte de l'information*. [En ligne] sur : <http://metiers.internet.gouv.fr/metier/architecte-de-l%E2%80%99information>. [Consulté le 22-03-2016].

MERCIER Sylvère [dit Silvae]. Bibliobsession. *La fonction ressource et les métiers de l'information-documentation*. [En ligne] sur : <http://www.bibliobsession.net/2010/06/02/la-fonction-ressource-et-les-metiers-de-linformation-documentation/> . [Consulté le 21-03-2016].

MEYRIAT Jean. Euroréférentiel-compétences et aptitudes des professionnels de l'I-D. *Association des professionnels de l'information et de la documentation (ADBS)*. Paris : 2004. [En ligne] sur : <http://certidoc.net/fr1/tout.php>. [Consulté le 22-03-2016].

MICHEL Jean. *Les professionnels de l'information-documentation à l'heure du document numérique et des réseaux électroniques*. [En ligne] sur : <http://michel.jean.free.fr/publi/JM293.html>. [Consulté le 21-03-2016].

POUCHOL Jérôme. Pratiques et politiques d'acquisition. *Bulletin des bibliothèques de France (BBF)*, 2006, n° 1, p. 5-17 [En ligne] sur : <http://bbf.enssib.fr/>. [Consulté le 21-03-2016]. http://www.journal.cybrarians.org/index.php?option=com_content&view=article&id=667

6-Chan, D (2005) Changing roles of reference librarians: The case of the HKUST Institutional Repository [online]. *Reference services review*, vol. 33, no. 3. Available on: < <http://repository.ust.hk/dspace/bitstream/1783.1/2039/2/p268.pdf> >(visited date 21/03/2016)

7-Chang, SH (2003) Institutional repositories: the library's new role [online]. *OCLC Systems & Services*, vol. 19, no. 3. Available on: _____ < <http://www.citeulike.org/user/chuy/article/11665681> >(visited date 21/03/2016)

8-Genoni, P (2004) Content in institutional repositories: a collection management issue [online]. *Library Management*, vol. 25, no. 6/7. Available on: <http://espace.library.curtin.edu.au/cgi-bin/espace.pdf?file=/2008/11/13/file_8/19655>(visited date 21/03/2016)

9-Jain, P(2006) *The Role of institutional repository in digital scholarly communications* [online]. Available on: < http://www.ais.up.ac.za/digi/docs/jain_paper.pdf >(visited date 21/03/2016)

10-Kamraninia1, Katayoon (2010) Librarians' role as change agents for institutional repositories: A case of Malaysian academic libraries [online]. *Malaysian Journal of Library & Information Science*, vol. 15, no.3. Available on: <

<http://ejum.fsktm.um.edu.my/article/961.pdf>>(visited date 21/03/2016)

11-Nixon, William (2002) The Evolution of an Institutional E-Prints Archive at the University Of Glasgow [online]. *Ariadne*, no. 32 . Available on:

< <http://www.ariadne.ac.uk/issue32/eprint-archives/>>(visited date 21/03/2016)

12-Phillips, H (2005) Leading roles for reference librarians in institutional repositories [online]. *Reference Services Review*, vol. 33, no. 3. Available on: < first.emeraldinsight.com>(visited date 21/03/2016)

13- Université Abou Bekr Belkaid Tlemcen (2016) *DSpace à L'université Abou Bekr Belkaid Tlemcen* [en ligne] . Disponible sur: < <http://dspace.univ-tlemcen.dz/>>(consulté le 21/03/2016)

14-University of Biskra (2016) *University of Biskra Repository* [online]. Available on: < <http://dspace.univ-biskra.dz:8080/jspui/>>(visited date 21/03/2016)

15- University of Biskra (2016) *University of Biskra Theses Repository*[online]. Available on: < <http://thesis.univ-biskra.dz/>>(visited date 21/03/2016)

16-Walters, Tyler (2007) Reinventing the library—How repositories are causing librarians to rethink their roles[online] . *Portal: Libraries and the Academy*, vol. 7, no. 2 . Available on:

<<https://smartech.gatech.edu/bitstream/handle/1853/14421/7.2walters.pdf;jsessionid=4C17C7A2887705914C0493AF2A828AD4.smart2?sequence=1>>(visited date 21/03/2016)

17-Ware, M (2006) *Pathfinder research on web-based repositories* [online]. Available on:<<http://mrkwr.files.wordpress.com/2006/11/pals-report-on-institutional-repositories.pdf>>(visited date 21/03/2016)

18- Webometrics (2016) *Arab World Ranking of Repositories* [online]. Available on: < repositories.webometrics.info/en/Arab_world >(visited date 21/03/2016)

Humanités numériques et métier d'archiviste : quelle articulation possible ?

Mohamed Lat Diop

Ebad Sénégal

Introduction

Un des phénomènes marquant de la deuxième moitié du XXe siècle et du début de ce siècle est l'irruption des Technologies de l'Information et de la Communication (TIC) dans tous les secteurs d'activité de l'Homme. Le domaine des Sciences de l'information et de la documentation, notamment ses disciplines telles que l'archivistique, la bibliothéconomie, l'édition, ou encore la documentation, ainsi que les métiers respectifs, n'ont pas échappé à ce mouvement qui se traduit aujourd'hui par une banalisation accélérée de l'utilisation des TIC, du numérique et du web dans leurs activités professionnelles.

Cette situation a conduit des chercheurs et des praticiens à questionner les usages des outils et médias sociaux et à analyser les transformations subies par ces domaines et leurs métiers, donnant ainsi naissance aux humanités numériques.

Le dernier numéro de la Revue Française de Sciences de l'information et de la communication de l'année 2016, coordonné par Julia Bonaccorsi, Valérie Carayol et Jean Claude Domenget et consacré à la rencontre entre les « Humanités numériques et Sciences de l'information et de la communication (SIC) »⁵² témoigne plus que jamais de l'actualité de la

⁵² Voir Jean-Claude Domenget, Julia Bonaccorsi et Valérie Carayol, « Introduction au dossier « Humanités numériques et SIC » », *Revue française des sciences de l'information et de la communication* [En

thématique des Humanités numériques et, notamment de leur articulation avec les SIC.⁵³ Dès lors, les chercheurs de cette discipline doivent s'interroger sur l'impact qu'ont les humanités numériques sur leurs champs et objets de recherche, de manière à savoir « *ce que les SIC ont à dire des HN tout autant que ce que les HN apportent à la réflexion en SIC* ». ⁵⁴

Pour ce faire, nous évoquerons d'abord les Humanités numériques, en retraçant leur histoire et en précisant leurs contours et leur Object. Ensuite, nous reviendrons sur les Sciences de l'information documentaire, en portant notre regard uniquement sur les archives, sa définition, l'étendu des documents qu'elles recouvrent, et les spécificités du métier d'archiviste, voire les transformations qu'il a subies. Enfin dans une troisième partie, nous verrons à quelle articulation est-il possible d'arriver entre les Humanités numériques et les archives.

Le choix est fait ici de ne traiter que les archives en rapport avec la problématique des Humanités numériques, leur articulation, tout en mettant de côté la bibliothéconomie et la documentation, disciplines par ailleurs connexes à l'archivistique. Ce choix est gratuit. Il se justifie par le fait que nous avons eu une formation initiale d'archiviste ; c'est donc notre spécialité. Autrement dit, nous nous sentons mieux à l'aise dans cette discipline que dans ses deux autres sœurs, à savoir la bibliothéconomie et la documentation, pour appréhender et saisir cette articulation. Aussi, cela nous met dans une position à même de nous permettre de mieux décrire ce que nous vivons ou observons chez les autres collègues archivistes ou enseignants d'archivistique.

Qu'est ce que les Humanités numériques ?

Les Humanités numériques ou *Digital Humanities* en anglais sont une fusion entre les Sciences Humaines et Sociales (SHS)

ligne], 8 | 2016, mis en ligne le 29 mars 2016, consulté le 30 mars 2016.
URL : <http://rfsic.revues.org/1984>

(Consulté le 31 mars 2016)

⁵³ Notons que dans ce propos, l'évocation des SIC fait référence aux Sciences de l'information documentaire, en particulier aux domaines des archives, de la documentation et des bibliothèques, considérés comme domaines à part entière de la grande famille des SIC.

⁵⁴ Jean-Claude Domenget, Julia Bonaccorsi et Valérie Carayol, op. cit.

et l'Informatique. Elles sont, plus clairement, l'application des technologies informatiques dans les méthodes de recherche en SHS. Roberto Busa est considéré comme l'un des pères fondateurs de cette discipline, avec son «Index Thomisticus» des années 1960.

Cependant, si la terminologie est relativement récente, la notion n'est pas facile à définir au regard de la diversité des points de vue et de la mouvance du concept.

Pour décrire ce que sont les Humanités numériques, nous nous référons aux termes du «Manifeste des Digital Humanities», résultats des journées d'étude ou THATCamp, organisé à Paris du 18 et 19 mai 2010 par les acteurs et observateurs des Humanités numériques, pour tenter de donner une idée claire de ce que c'est que les Humanités numériques. La Manifeste donne en effet trois éléments de définition importants.

Le premier pointe le doigt sur le *«tournant numérique pris par la Science et qui modifie et interroge les conditions de production et de diffusion des savoirs.»*⁵⁵

Le deuxième point fixe les territoires des Humanités Numériques en précisant qu'elles concernent uniquement les SHS, les Arts et les Lettres. Les H. N. ne sont pas amnésiques de leur passé : *«elles s'appuient, au contraire, sur l'ensemble des paradigmes, savoir-faire et connaissances propres à ces disciplines, tout en mobilisant les outils et les perspectives singulières du champ du numérique.»*

Le troisième et dernier élément de définition de la Manifeste insiste sur le caractère multilingue et multidisciplinaire du mouvement des H. N., un mouvement né dans les disciplines des SHS, en réaction au phénomène du *«savoir lire numérique»*⁵⁶.

En résumé, les Humanités numériques sont une nouvelle discipline dédiée aux technologies numériques utilisées par les humanités dans la production de savoirs. Les pratiques de recherche en SHS sont donc au croisement des technologies numériques et des humanités.

⁵⁵ THATCamp. Manifeste des Digital Humanities : <http://tcp.hypotheses.org/318> consulté le 27/ 03/2016

⁵⁶ Doueïhi, Milad. *Savoir Lire numérique*. In : *La Grande conversion numérique*. Paris, Le Seuil, 2008, p 13.

Elles étudient notamment l'impact de ces techniques sur le patrimoine culturel, et les institutions en charge de la mémoire, les Bibliothèques, les Archives ainsi que cette nouvelle dimension qu'est et la culture numérique.

Les acteurs des Humanités Numériques sont conscients du fait que la recherche en SHS est de plus en plus inséparable des outils développés par les nouvelles technologies. Ces outils et services constituent des réservoirs de données numériques, des centres, de réseaux et de dispositifs informatiques puissants utiles aux chercheurs bref, des cyber-infrastructures, qui permettent le développement de cette nouvelle discipline.

Le terme cyber-infrastructure a été évoqué pour la première fois par la National Science Foundation (NSF)⁵⁷ pour décrire «*les nouveaux environnements de recherche dans lesquels les capacités les plus abouties des outils informatiques seraient à la disposition des chercheurs dans le cadre d'un réseau interopérable.*»⁵⁸ Il a été défini pour la première dans le rapport publié par le Conseil américain des sociétés savantes en 2006 et qui est considéré comme le texte fondateur des Humanités numériques. Le rapport définit la notion de cyber-infrastructure comme «*une couche d'informations, d'expertises, de normes, de politiques, d'outils et de services qui sont partagés par les communautés de recherche, mais développés à des fins scientifiques spécifiques: la cyberinfrastructure est quelque chose de plus spécifique que le réseau lui-même, mais elle est quelque chose de plus général qu'un outil ou qu'une ressource développé pour un projet particulier, un ensemble de projets, ou, plus largement encore, pour une discipline particulière.*»

Les Humanités numériques sont donc une «*nouvelle façon de faire de la recherche en SHS. C'est la construction d'un web de données scientifiques dans le web de données, qui, lui, est une extension du web actuel où les données numérique sont structurées et liées entre elles.*»⁵⁹ Les acteurs de ce mouvement

⁵⁷ Fondation Nationale de la Science (NSF) <http://www.nsf.gov/#1> (consulté le 31 mars 2016)

⁵⁸ Voir <http://dhi.intd.cnam.fr/digital-humanities/> (consulté le 31 mars 2016)

⁵⁹ Auparavant s'appelait « Consortium pour ordinateurs en sciences humaines / Consortium for Computers in the Humanities (COSH/COCH) », Voir <http://dhi.intd.cnam.fr/digital-humanities/> (consulté le 31 mars 2016)

ont estimé que le patrimoine des SHS reste encore largement inexploité. Ils considèrent que seules les productions des chercheurs, autrement dit les publications scientifiques, ont suscité des questionnements et réalisations numériques (HALSHS, Cairn, Revues.org) qui sont de nouvelles pratiques d'édition électronique, et que les initiatives lancées concernant les fonds documentaires sont peu nombreuses.

Ces fonds sont des archives conservées dans les bureaux des chercheurs ou dans des disques durs, dans les centres de documentation ou bibliothèques de recherche. Ils sont aujourd'hui les principales sources de recherche des chercheurs. Ils ont une dimension patrimoniale et constitueront une vraie richesse à long terme pour les institutions de recherche qui les auront numérisés et diffusés. Mais en plus de ces fonds documentaires anciens, existent aujourd'hui les données nativement numériques appelés également archives électroniques (photos, textes, e-mails, données 3D, Bases de données, sons, flux RSS, sites web, etc.)

Ces fonds de données, nativement numériques, de plus en plus importants, et cela dans de très nombreuses disciplines en SHS, doivent leur existence et leur accessibilité aux outils et services numériques qui les ont générés. La rapide obsolescence de ces outils et de leur format est cependant une réalité qui peut entraîner l'altération ou la perte de ces données.

Les acteurs des Humanités numériques décident alors de les faire rentrer dans l'extension du web, devenu aujourd'hui un web de données ou Big Data.

Si l'objectif général des Humanités numériques est de diffuser les données de la recherche en SHS, les objectifs spécifiques quant à eux sont de trois ordres :

- La diffusion massive et l'édition de presque de tous les matériaux utilisés par le chercheur et l'enseignant ainsi que les données de la recherche,
- La démultiplication des accès aux documents et à l'information tout en assurant l'interopérabilité des données,
- La structuration de l'information contenue dans ces documents.

Les pays anglo-saxons sont les précurseurs des Humanités numériques et ils ont créé les centres d'Humanités Numériques les plus visibles. Aux Etats-Unis, la création de l'«Institute for Advanced Technology in the Humanities» (IATH) en 1992, marque le début de la concrétisation du concept d'Humanités numériques. Par la suite sont apparues des structures comme le centre «Roy Rosenzweig» de l'université de George Mason et il faut souligner que «Zotero», un outil d'aide au référencement bibliographique et «Omeka», pour la publication en ligne, sont hébergés à Washington. L'Angleterre a suivi, avec le «e-Research Center» de Oxford, fondé en 2006, et une dimension académique par l'introduction d'un cursus de master et de doctorat en Humanités numériques au University College London (UCL) et au King's College. L'Australie et le Canada aussi ont initié des expériences d'Humanités Numériques. Le Canada avec la Société Canadienne des humanités numériques⁶⁰ créée en 1986, et l'Australie avec l'Australasian Association for Digital Humanities (aaDH), créée en mars 2011.⁶¹

Après ce bref historique, nous allons maintenant aborder les archives, le métier d'archiviste, ses spécificités et les transformations qu'il a subies à l'ère du numérique, du Web et des TIC, avant de voir leur articulation avec les Humanités numériques.

Les archives, le métier d'archiviste, ses spécificités et ses évolutions

Tous les textes législatifs et réglementaires de presque tous les pays, définissent les archives comme étant l'ensemble des documents quels qu'en soient la nature, la forme et le support produit ou reçu par un organisme public ou privé dans l'exercice de ces activités. Cependant, elles peuvent avoir deux statuts : public ou privé.

Les archives sont considérées comme publiques quand elles proviennent de l'activité de l'Etat, de l'administration centrale et des services extérieurs de l'Etat régionaux et départementaux ou des collectivités territoriales (région, département, communes). Toutes ces entités sont des producteurs d'archives publiques. Cependant, celles-ci peuvent également provenir

⁶⁰ Voir <http://csdh-schn.org/> consulté le 31 juillet 2016

⁶¹ Association Australienne des humanités numériques, voir <http://aa-dh.org/> consulté le 31 juillet 2016

d'établissements publics, ou d'organismes ou personnes de droit public et exerçant une activité spécifique sous la tutelle de l'Etat.⁶²

Les archives des entreprises publiques, et sociétés nationalisées ainsi que celles produites par des organismes de droit privé, mais chargés d'une mission de service public, et les minutes et répertoires des officiers publics et ministériels (notaires), sont également des archives publiques à l'exception des dossiers des clients des notaires qui ont un caractère privé.

Globalement on peut dire que les archives publiques le sont par essence, par naissance ou par propriété. Elles sont imprescriptibles, c'est à dire leur statut ne pourra subir aucune atteinte avec le temps, il sera toujours valable car elles resteront éternellement publiques ; inaliénables, c'est à dire que les archives ne peuvent être aliénées, cédées ou vendues; et revendicables, c'est à dire elles peuvent être réclamées par l'Etat.

Les archives privées quant à elles sont toutes les archives qui ne rentrent pas dans le cadre des archives publiques. Elles peuvent donc être définies en creux de celles-ci.

Les spécificités du métier d'archiviste et leurs conséquences

Le Dictionnaire de terminologie archivistique de la Direction des Archives de France (DAF) définit l'archivistique comme étant la « Science qui étudie les principes et les méthodes appliquées à la *collecte*, au *traitement*, à la *conservation*, à la *communication* et à la mise en valeur des *documents d'archives* » et l'archiviste comme étant « le professionnel chargé de la gestion des archives »⁶³.

Le métier d'archiviste se distingue des autres métiers du patrimoine type bibliothécaire, documentaliste, par au moins trois spécificités :

La première concerne la notion de fonds d'archives. Une des grandes différences avec les métiers de bibliothécaire et de documentaliste repose sur la notion de fonds conçu comme un

⁶² Cf. Les établissements publics hospitaliers, des établissements scolaires et les organismes consulaires (chambre de commerce, chambre des métiers, Chambre de l'Agriculture).

⁶³ DAF (2002), Dictionnaire de terminologie archivistique, p10-11

ensemble entier, qui se construit sans l'archiviste. Le fonds d'archives est généré par le producteur d'information et est issue de son activité. L'archiviste ne le choisit pas.

Les archives sont ainsi intimement liées au producteur, et à son histoire, et sa connaissance est nécessaire à la celle du fonds.

La deuxième spécificité des archives est qu'elles sont un patrimoine proche des métiers des musées. Les archives sont des pièces uniques et originales, et elles ont un caractère précieux, conféré par leur unicité et leur ancienneté. Elles sont le témoignage et la première source de l'histoire.

En plus de leur caractère historique, les archives ont également un caractère « humain » et « personnalisé » car elles touchent les hommes et les femmes au plus près. En effet, les archives sont le seul domaine du patrimoine qui touche la vie privée des individus.

Enfin, les archives peuvent avoir une valeur juridique, surtout celles qui sont les fondements et la base de la société, tel que l'Etat civil qui prouve notre existence ou le Cadastre qui est la base de la propriété.

Ces trois spécificités des archives ont des incidences sur le métier de celui qui l'exerce, à savoir l'archiviste. Ainsi pour mettre en œuvre le principe du respect des fonds⁶⁴, considéré comme l' « *assise théorique, la loi qui régie toutes les interventions archivistiques* », ⁶⁵ la connaissance du service producteur est nécessaire pour le traitement et l'utilisation des archives. Par ailleurs la conservation et la destruction sont réglementées et obéissent à des règles strictes.

Sur le plan patrimonial et juridique, l'archiviste a l'obligation de garder les archives qui font preuve. Ne pas les conserver, c'est compromettre, voire priver, leur propriétaire ou la société en général, de leurs bases juridiques

⁶⁴ Le Dictionnaire de terminologie archivistique définit ainsi le principe de respect des fonds comme étant le « principe fondamental selon lequel les archives d'une même provenance ne doivent pas être mélangés à celles d'une autre provenance ; ce principe a parfois inclus le principe de respect de l'ordre primitif.» Le Conseil International des Archives de son côté l'a défini en 1964 comme étant le « principe selon lequel chaque document doit être placé dans le fonds dont il provient t, dans ce fonds, à sa place d'origine.»

⁶⁵ Carol, C. et al. 2003 Fondements de la discipline archivistique, p 64

Sur le plan historique, il existe un facteur de risque non négligeable, car les archives peuvent être détournées pour des raisons idéologiques. L'archiviste est donc celui qui veille à la bonne conservation des documents qui ont un caractère et un enjeu historique. Le caractère parfois précieux de certains documents fait que les archives courent parfois des dangers (vols, éparpillement, destruction volontaire). Enfin les archives peuvent également avoir une valeur marchande, et elles peuvent donc être volées.

Les archives ont un rapport avec la vie privée concernant la communicabilité de certains documents, ce qui fait peser une grande responsabilité sur l'archiviste. C'est pourquoi, les archives qui touchent à la vie privée ne sont pas communicables avant des délais spéciaux. Or, avec le web social, la question de la protection des données à caractère personnel est plus que jamais posée. Si la sensibilité de certains documents d'archives est la raison de l'existence de ces délais spéciaux pour leur communication, le principe même de l'Internet et du Web 2.0 sont en contradiction totale avec cette pratique.

De ce fait, l'archiviste se doit d'être vigilant et prudent quant à la communication, en ligne, ou hors ligne, de certains documents confidentiels ou d'informations à caractère personnel.

Les problèmes de l'accès aux archives font référence au fait que celles-ci sont des sources et documents bruts. Leur bonne utilisation pose des problèmes juridiques et réglementaires. L'archiviste doit être celui qui facilite l'accès aux documents à l'utilisateur, en les décrivant et en les analysant afin de les rendre intelligibles et contextualisés. Il est aussi censé avoir des connaissances juridiques et réglementaires relatives à l'accès et la communicabilité des documents, mais aussi savoir reconnaître l'original de la copie en termes de diplomatique contemporaine.

Les problèmes auxquels l'archiviste est désormais confronté

Nous pouvons identifier au moins trois problèmes relatifs aux archives et à l'exercice du métier d'archiviste à savoir la production en masse de documents et d'informations, la question des nouveaux supports, ainsi que l'exigence de transparence et les besoins d'information du « nouveau type » de citoyen.

Concernant la production en masse de l'information, l'informatique et les réseaux numériques ont fini d'augmenter la production de documents, et donc des archives. L'archiviste se doit, pour faire face à cette inflation documentaire, d'agir au niveau des services producteurs et des secrétariats, leur donner des outils de gestion, préparer le travail d'archivage en amont avec des tableaux de gestion, et des bordereaux de versement, qui sont les principaux leviers sur lesquels il peut s'appuyer pour faire face à ces transformations, apportées par les nouveaux outils de production de document.

Ensuite, les nouveaux supports des documents ont également fait évoluer le métier d'archiviste qui doit désormais allier la gestion des nouveaux supports à celle des supports traditionnels. Ces supports numériques et/ou électroniques posent des problèmes de conservation et de pérennisation que les normes ICA Req, Moreq, La NF Z 42 013 et l'OAIS tentent de corriger en apportant des solutions de conservation et de stockage pérennes. Ces normes et standards suggèrent globalement que l'archivage numérique est différent de l'utilisation de l'informatique dans les archives⁶⁶, et que celui-ci est également différent de la sauvegarde informatique.

De plus, l'Internet a également impacté le métier d'archiviste par l'entrée des réseaux sociaux numériques dans les territoires des Sciences de l'information documentaire, axe que nous développons plus loin.

Enfin les archives sonores et audiovisuelles, qui sont également des documents périssables, ont la spécificité de ne pas être lisibles sans l'utilisation d'un appareil de lecture ou d'une machine. En raison de la durée de vie limitée des supports (5-10 ans), leur conservation peut nécessiter parfois une migration qui peut avoir un coût financier énorme.

Quant à l'exigence de transparence du citoyen, il est maintenant admis que les besoins du public ont évolué. L'exercice de la citoyenneté, parallèlement à la diminution de la recherche historique qui n'est plus un domaine porteur, influe sur la pratique professionnelle de l'archiviste qui doit désormais

⁶⁶ Françoise Banat-Berger, Laurent Duploux et Claude Huc. (2009). L'archivage numérique à long terme : les débuts de la maturité : Manuel et guide pratique. Paris : La documentation française.

répondre aux besoins d'informations des citoyens, et ce dans la diligence. Le citoyen a changé de comportement informationnel. Il est de plus en plus exigeant, et demande un accès rapide, et simple à l'information.

L'accès aux documents administratifs est maintenant garanti par des organismes comme la Commission d'Accès aux Documents Administratifs (CADA) en France ou au Sénégal la loi du 30 juin 2006 relative aux archives et aux documents administratifs. Le citoyen peut désormais exiger de consulter un document qui le concerne ou pas, à condition que la communication des documents ne porte atteinte à la sûreté publique ou à la vie privée des individus.

Le principe est que toutes les archives sont désormais communicables à quelques exceptions, avec des délais spéciaux. Les usagers ont ainsi droit à des copies et déjà les généalogistes revendiquent des copies d'actes d'Etat civil.

Partant de toutes ces considérations, nous allons maintenant nous intéresser aux articulations possibles entre le monde des archives et les Humanités numériques.

Articulation entre archives et Humanités numériques

Par Humanités numériques pour les archives, nous faisons d'abord référence à l'utilisation des outils et médias numériques en appuie aux archives et au métier de l'archiviste. Autrement dit, nous nous plaçons dans la problématique de la « médiatisation » ou de la « technologisation » de cette discipline et du métier, grâce aux TIC, au Web et aux réseaux numériques, qu'ils soient sociaux, associatifs ou professionnels. Cette médiatisation se caractérise par la présence de « médias » et ou « machines à archiver » et à diffuser dans les archives. Elle peut donc aller de la simple utilisation d'un ordinateur ou d'un logiciel d'archivage, ou même de bureautique, dans le cadre de la chaîne de traitement archivistique, à la mise en place d'un projet de numérisation, de mise en ligne d'instruments de recherche ou d'un vrai Système d'Archivage Electronique (SAE), voire de Gestion Electronique des Documents (GED).

La GED se définit comme étant l'utilisation de moyens informatisés pour l'ensemble de la gestion d'un document numérique (fichiers texte, fichier tableur, image, vidéo, fichier audio etc.) Elle « *recouvre tous les processus visant à gérer et organiser des documents de manière informatisée. Elle touche*

*toutes les étapes du cycle de vie d'un document, de l'acquisition à l'archivage», de l'acquisition à l'archivage en passant par le classement, le stockage et la diffusion.*⁶⁷

Quant au SAE, c'est un système de gestion des archives numériques et/ou électroniques, avec différentes fonctionnalités visant à garantir les caractéristiques suivantes du « records » ou archives qu'il gère : l'authenticité, la pérennité, la fiabilité et l'exploitabilité.

Plus précisément, la dimension documentaire et archivistique des Humanités numériques peut s'exprimer de diverses manières : communication de documents dans et hors des institutions patrimoniales type archives, bibliothèques et centres de documentation, formation des professionnels de l'information documentaire,⁶⁸ et enfin publicité et valorisation de la profession et des métiers relatifs aux archives et aux bibliothèques.

En ce sens, la perspective des Humanités numériques peut se montrer pertinente pour penser le monde des archives et des bibliothèques, notre rapport à l'information documentaire, au document d'archives, au livre, ainsi que les modalités de leur gestion, de leur traitement, et de leur accès. « *Les Humanités numériques se fixent pour objectif de favoriser le développement d'une culture et d'une pratique technologique en SHS. En tant que champ de recherche émergent, les HN s'intéressent aux technologies du point de vue des usages ou dans une perspective*

⁶⁷ Voir <http://ged.prestataires.com/conseils/ged> consulté le 31 juillet 2016

⁶⁸ On peut prendre l'exemple du MOOC « *Bien archiver : la réponse au désordre numérique* » qui en est en sa deuxième session. On peut lire sur la page dédiée au MOOC sur site de France Université Numérique (FUN) que : « le cours a pour objectif d'expliquer les concepts essentiels de l'archivage managérial : analyse des besoins de preuve et de mémoire, qualification des données, élaboration de règles de vie des documents et des règles du jeu pour tous les acteurs concernés. » Selon Marie Anne Chabin, sur le site de son cabinet Archives17, il est « incontournable pour toute personne impliquée dans un projet d'archivage. » « C'est le premier MOOC en français sur l'archivage et il a l'avantage de présenter deux parcours pour deux publics : 1. un parcours « Découverte » (3 h en tout) qui est une sensibilisation des acteurs de l'entreprise (décideurs et collaborateurs) à l'importance des traces numériques, conscientes ou non, produites au quotidien au moyen des outils numériques et aux enjeux liés au temps ; 2. et un parcours à destination des « Chefs de projet » (documentalistes, archivistes ou informaticiens). L'ensemble est accompagné d'une étude de cas qui se déroule sur cinq des six semaines. »

*historique. Elles posent également la question de la conception de ces technologies et du rapport critique qu'il convient d'entretenir avec elles.»*⁶⁹

Bien qu'existant, les liens entre la recherche dans le domaine des HN et celle portant sur les archives et bibliothèques sont encore peu formalisées. C'est pourquoi nous nous proposons d'explorer les différentes formes que pourraient prendre ces liens, qu'il s'agisse de projets de numérisation d'archives ou de projets de records management, de bibliothèques numériques ou virtuelles, ou encore de sites ou blogs de professionnels dédiés aux archives.

Le Web contient beaucoup de documents mais est relativement vide de «données brutes» et d'informations structurées et l'accessibilité de ces documents est assurée depuis les années 90 par les moteurs de recherche type Altavista ou Google. Dès lors, le mouvement des Humanités numérique a décidé de se joindre à celui du libre accès, pour construire des espaces de données plus larges, mondiaux, afin d'offrir aux scientifiques des corpus de données numériques, documentés, accessibles, interopérables et pérennes.

Les corpus de données numériques que sont les bibliothèques numériques ou virtuelles, la numérisation des sources, le web sémantique, l'open data, la cyber-infrastructure, etc., sont la traduction concrète des Humanités numériques.

En France, on ne compte plus le nombre de projets de numérisation et de mise en ligne d'archives, de documents d'Etat civil ou généalogiques. La SIV⁷⁰ (Salle des Inventaires Virtuelle) des Archives Nationales de France est là pour nous dire que plus d'un million d'archives papier sont numérisées par la Direction des Archives de France, avec des inventaires les décrivant et des notices décrivant les producteurs de ces documents. Nombreuses aussi sont les Archives Départementales en France ayant numérisé et mis en ligne leurs instruments de recherche.

Au Sénégal, la numérisation des fonds d'archives des institutions

⁶⁹ Voir <http://celluloid.hypotheses.org/929> (consulté le 31 mars 2016)

⁷⁰ Voir <https://www.siv.archives-nationales.culture.gouv.fr/siv/cms/content/display.action?uuid=Accueil1Root&onglet=1>

et des archives de l'AOF⁷¹ conservées aux Archives Nationales, ainsi que le grand projet de numérisation « *des patrimoines et fonds documentaires audiovisuels, iconographes, sonores et textuels de l'Institut fondamental d'Afrique Noire Cheikh Anta Diop* », ⁷² sont les premiers pas vers le mouvement des Humanités numériques.

Les fonds documentaires et les données sources pour la recherche en SHS ont désormais pris le tournant du numérique selon Stéphane Pouyllan⁷³, ouvrant ainsi la voie à l'ère des usages numériques en SHS. La publication simultanée et quasi-instantanée de résultats de la recherche via un blog ou une archive ouverte constitue également une avancée relevant des Humanités numériques.

A titre d'exemple, on peut citer le blog d'Olivier Le Deuff de Olivier Le Deuff, chercheur en Sciences de l'information et de la communication (SIC), intitulé « Le Guide des égarés ». Ce blog est une vitrine de ces activités d'enseignement et de recherche et traite notamment de « cultures digitales, culture de l'information, Information literacy, documentation, CDI, Pédagogie, sciences de l'information et de la communication, Humanisme numérique, Humanités digitales. »⁷⁴

Sur son blog, ce chercheur publie toutes ses publications (chapitres d'ouvrage, articles scientifiques, actes de colloque ou de conférence, billets etc.). Il dépose également ces publications sur l'archive ouverte Hyper Article en Ligne (HAL) et les liens vers ces publications déposées sur HAL sont également disponibles à partir de son blog.

On peut également citer l'exemple du blog de Mor Dièye, enseignant-chercheur en archivistique à l'EBAD, pour qui son

⁷¹ Voir <http://fr.allafrica.com/stories/201404170911.html> (consulté le 31 mars 2016)

⁷² Voir <http://fr.allafrica.com/stories/201004080548.html> (consulté le 31 mars 2016)

⁷³ POUYLLAN, Stéphane. Construire le web de données pour les SHS fonds documentaires scientifiques de la recherche en SHS. P 5 Disponible ici: http://archivesic.ccsd.cnrs.fr/docs/00/51/68/69/PDF/WebofData_CNRS_SHS_ElementsADONIS_CN2SV_2010_v2-4.pdf consulté le 20/03/2016

⁷⁴ Voir : http://www.guidedesegares.info/2015/10/11/fadben-2015-la-peche/?utm_source=dlvr.it&utm_medium=twitter (consulté le 29 mars 2016)

site est un « un espace d'expression et d'échanges. »⁷⁵ Il y publie des articles scientifiques, résultats de ses recherches en archivistique, des commentaires sur l'actualité d'une manière générale et en rapport avec le monde des archives et le métier d'archiviste en particulier. Le lien vers sa thèse déposée sur HAL est également disponible à partir de son blog.

« Transarchivistique »⁷⁶ est un autre blog, tenu par Anne Marie Chabin, une archiviste-paléographe, diplômée de l'Ecole des chartes, qui, après avoir été professeure au Conservatoire National des Arts et Métiers en France (CNAM) de Paris (France), est aujourd'hui, experte en stratégie et méthodologie d'archivage et conseillère en Records managements et archivage électronique à travers son cabinet Archives17.⁷⁷ Elle définit son blog comme un « *blog d'humeur destiné aux praticiens des archives et de l'archivage.* »

Pour revenir sur le cas du Sénégal, Atoumane Mbaye, archiviste aux Archives Nationales du Sénégal, tient un blog qui se veut relater l'histoire du Sénégal de 1455 à nos jours, avec des archives sous forme de fichiers audio, d'éphémérides, de chaque jour, une présentation des rois et présidents du Sénégal, des archives portant sur les personnages et lieux du Sénégal, etc.

Nous citons l'auteur de ce site pour qui ce travail « *se veut comme une contribution à la sauvegarde de la mémoire collective de la nation sénégalaise à travers une plateforme de recherche des dates marquantes de l'histoire du Sénégal, de la période des comptoirs européens à nos jours. Ressource de premier plan pour les historiens et chercheurs en sciences sociales, le site vise un grand public qui va de l'universitaire au simple curieux en mal de repères chronologiques sur l'histoire du pays.* »⁷⁸

Ces sites ou blogs sont des preuves de l'appropriation des outils numériques et du Web par certains chercheurs en Sciences de l'information documentaire en général et en archivistique en particulier.

Sur un autre plan, l'ouverture des archives audiovisuelles de la

⁷⁵ Voir <http://www.mdieye.com/> (consulté le 31 mars 2016)

⁷⁶ Voir <http://transarchivistique.fr/> (consulté le 31 mars 2016)

⁷⁷ Voir <http://www.archive17.fr/index.php/archiver-avec-marie-anne-chabin/> ou <http://www.archive17.fr/>

⁷⁸ Voir <http://www.histoire-senegal.com/> (consulté le 31 mars 2016)

Croix Rouge Internationale sur Internet à travers une plateforme⁷⁹ est une autre preuve récente de la prise en compte des opportunités offertes par le numérique par du monde des archives.

Ces démarches témoignent de l'influence du numérique et des outils technologiques sur le monde des archives et sur le métier d'archiviste, mais aussi et surtout de la transformation en cours du métier de « chercheur en archivistique » et de la « science » ou discipline archivistique même, de plus en plus en phases avec les technologies numériques.

Dans un autre registre et sur un autre ton, on peut citer les blogs et Tumblr⁸⁰ comme -Archives et cultures pop,⁸¹ qui sont dédiés aux Archives. Selon l'auteur du Tumblr « Archives et cultures pop », cet outil est une manière de « *découvrir le monde des archives à travers la littérature populaire, les séries et le cinéma* ».

Dans la même perspective, Déborah Pidet, jeune archiviste emprunte le même ton pour parler de son quotidien d'archiviste, de records manager dans une grande entreprise française et de chargée de cours au Département Archives Médiathèque de l'Université Jean Jaurès, ex Toulouse 2 Le Mirail. Elle parle d' « *l'archivage hybride, moitié papier, moitié numérique* » dans son blog dédié, toujours selon elle, aux « *tribulations d'une archiviste en Archivie : Racontars archivistiques ... et autres !* »⁸²

A travers ces quelques exemples, nous voulons montrer comment certains archivistes, enseignant (e) s et chercheur (e) s en archivistique se sont appropriés le numérique et les opportunités qu'il offre, pour valoriser leur travail, des fonds

⁷⁹ Voir <https://avarchives.icrc.org/> (consulté le 31 mars 2016)

⁸⁰ Un Tumblr est, selon Wikipédia, une plate-forme de microbloguement créée en 2007 et permettant à l'utilisateur de poster du texte, des images, des vidéos, des liens et des sons sur son tumblelog. Elle s'appuie principalement sur le reblogage.

⁸¹ Voir <https://archivespop.wordpress.com/> ou http://blog-phoenix-noir.tumblr.com/post/141023040829/better-call-archivist?utm_source=dlvr.it&utm_medium=facebook (consulté le 31 mars 2016)

⁸² Voir <https://archivisteenarchivie.wordpress.com/author/deborarchives/> (consulté le 31 mars 2016)

d'archives, ou partager leurs publications et travaux portant sur les archives.

Retenons également que tous ces sites et blogs se retrouvent parallèlement sur les réseaux sociaux type Facebook ou Twitter sous forme de Fans page, de profils, ou encore de comptes Twitter, accessibles à partir du blog ou sites en question. Cela leur donne un gain de popularité auprès d'un public peu initié en général, et qui ne ferait pas nécessairement la démarche d'aller visiter ces sites ou blogs spécialisés, d'où leur impact social.

Ces quelques exemples sont loin d'être isolés et notre évocation des cas de couplage entre Archives et numérique est également loin d'avoir fait le tour de la question.

La notion d'Humanités numériques recouvre également un autre entendement en plus de celui de «cyber-infrastructure», il s'agit de la notion d'«e-Science».

Les sens de ces deux notions importantes se croisent dans le fait que toutes les deux *«font référence à l'utilisation d'outils informatiques puissants et en réseau permettant de nouvelles collaborations scientifiques et des méthodes inédites de recherche.»*

L'e-Science est une notion récente, née au Royaume Unis à la fin du XX^e siècle, de John Taylor et qui signifie une nouvelle façon de faire de la science. Concrètement, les américains l'assimilent à la notion de cyber-infrastructure. Les anglais, et plus généralement les européens parlent d'E-Science à la place de cyber-infrastructure pour désigner le partage de ressources afin de permettre de nouvelles expériences scientifiques et de motiver de nouvelles collaborations. Elle est définie comme *«un nouvel environnement de recherche dans lequel des outils informatiques puissants sont mis à disposition des chercheurs au sein d'un réseau interopérable.»* L'e-Science permet l'optimisation des infrastructures pour le partage et l'exploitation des résultats de la recherche.

Une collaboration mondiale entre les grandes disciplines de la science est aujourd'hui possible grâce à l'e-Science, mais des infrastructures puissantes, interopérables et en sécurité sont la garantie d'un tel idéal : *«les cyberinfrastructures favorisent donc les communautés multidisciplinaires, les espaces de travail collaboratif, un peer-review plus immédiat et interactif. En effet, elles permettent de faire tomber les frontières entre les disciplines (grâce à un souci d'interopérabilité) et de nouvelles*

collaborations. C'est donc l'idée d'une « big science » et d'équipes de recherche virtuelles permises par des infrastructures informatiques puissantes.»

L'évocation de l'e-Science recouvre donc deux entendements : Internet et la science. Un nouveau couple est ainsi créé. C'est l'ère d'une nouvelle science ou la science qui renaît, sous une autre forme, une science à grande échelle, permettant une collaboration entre scientifiques situés partout dans le monde : *«une science collaborative. e-Science permet donc aux scientifiques de faire de la science autrement.»*

Ce changement de paradigme s'accompagne d'une nouvelle approche de la science, c'est la science démocratique ou la *«démocratisation de la science.»* Une science gratuite et accessible à tous à l'image de l'Internet, une science dont les notions de partage, de participation et de collaboration constituent le fondement théorique et sont le principe absolu.

Retenons simplement pour conclure cette réflexion, que les sens de ces deux notions importantes se croisent dans le fait que toutes les deux *«font référence à l'utilisation d'outils informatiques puissants et en réseau permettant de nouvelles collaborations scientifiques et des méthodes inédites de recherche.»*

Conclusion

La rencontre entre les SHS et l'Informatique a donné naissance au concept d'Humanités numériques, né aux Etats Unis au début des années 1990, et qui s'est généralisé à partir des années 2010, par ailleurs date du Manifeste des Digital Humanities. La création de plusieurs associations ou centres d'Humanités numériques dans plusieurs pays, notamment anglo-saxons, atteste de l'émergence de ce mouvement.

Si la rencontre entre les Humanités numériques et les Sciences de l'information et de la communication a déjà été théorisée et fait l'objet de travaux de recherche, il y a également la nécessité de mener la réflexion sur son articulation avec les disciplines telles que l'archivistique, la bibliothéconomie et la documentation. C'est à un tel exercice que nous avons essayé de nous livrer dans cet article, pour tenter d'appréhender et de saisir les rapports et les relations possibles et observables entre ce mouvement des Humanités numériques et le monde des archives notamment.

Dans la pratique, notre analyse a montré que les exemples sont nombreux, de concrétisation du concept d'Humanités numériques dans le champ des archives. Nous avons essayé de montrer comment elles contribuent à mettre en valeur ou promouvoir la discipline et le métier d'archiviste, tout en garantissant le respect des principes qui les fondent. Nous avons également essayé de donner à voir des exemples de valorisation du patrimoine documentaire et archivistique, par les outils numériques et du Web.

En perspective, il serait intéressant de tenter de mener la réflexion sur les rapports entre les Humanités numériques et les autres disciplines de la grande famille de l'info-doc, à savoir la bibliothéconomie, la documentation, ou encore l'édition.

Références

BANAT-BERGER Françoise, DUPLOUY Laurent et HUC Claude. (2009). *L'archivage numérique à long terme : les débuts de la maturité : Manuel et guide pratique*. Paris : La documentation française.

BIGOT Jean-Edouard, JULLIARD Virginie et MABI Clément, « Humanités numériques et analyse des controverses au regard des SIC », *Revue française des sciences de l'information et de la communication* [En ligne], 8 | 2016, mis en ligne le 24 mars 2016, consulté le 31 mars 2016. URL : <http://rfsic.revues.org/1783>

DOUEIHI, Milad. *Savoir Lire numérique*. In : *La Grande conversion numérique*. Paris, Le Seuil, 2008, p 13

DOMENGET Jean-Claude, BONACCORSI Julia et CARAYOL Valérie, « Introduction au dossier « Humanités numériques et SIC » », *Revue française des sciences de l'information et de la communication* [En ligne], 8 | 2016, mis en ligne le 29 mars 2016, consulté le 30 mars 2016. URL : <http://rfsic.revues.org/1984>

LE DEUFF Olivier (dir.) (2014), *Les temps des humanités digitales. La mutation des sciences humaines et sociales*,

Limoges, FYP éditions.

LE DEUFF, Olivier. (2015) Les humanités digitales précèdent-elle le numérique ?. Imad Saleh. *H2PTM 15*, Iste éditions, 978-1784051150. <sic_01220978>

MOUNIER Pierre (dir.), *Read/Write Book 2 : Une introduction aux humanités numériques*, Marseille, Open Edition Press, 2012, [En ligne] <http://books.openedition.org/oep/226>

THATCamp. Manifeste des Digitals Humanities : <http://tcp.hypotheses.org/318> consulté le 27/ 08/2012

Les Archives à l'ère des Humanités Numériques : Quelles mutations ?

Fatma Bilel & Abderrazak Mkadmi

Unité de recherche BNP, ISD, Université de la Manouba

fatma_bilel@yahoo.fr,

amkadmi@gmail.com

Introduction

« *Est-il possible qu'une chose soit à la fois numérique et humaniste ?* » (Scholes 2008) Voilà ce qu'on peut se demander lorsqu'il s'agit de penser au futur des sciences humaines et sociales (SHS) à l'ère du numérique, à ce couplage entre l'humain et le numérique qui paraît au début incompréhensible et ambigu s'ajoute la question de l'articulation entre les archives : mémoire humaine et le numérique comme "nouveau genre de mémoire" numérique.

Contrairement aux autres sciences – de la vie de la matière et de la vie- les SHS étaient considérées comme des disciplines qui "traitent de l'humain, des organisations, des échanges, de la communication" (Djamel, 2013), elles sont caractérisées par leur complexité liée à la structure et à la nature des êtres humains. Depuis longtemps on pouvait penser que les SHS n'ont rien à voir avec les technologies numériques. Seules les sciences exactes peuvent être mises sur le compte du numérique, or, il paraît que cette idée est fautive étant donné que la littérature est paradoxalement, parmi tous les arts, a rencontré l'informatique le plus tôt (Clément, 2001). Avec l'arrivée de l'ordinateur dans la sphère des SHS, et avec le développement spectaculaire des nouvelles technologies de l'information et de la communication moyennant le web, les SHS ont vécu un bouleversement profond majeur touchant ses méthodes et ses paradigmes.

Dès lors, et avec l'introduction des outils informatiques dans les recherches humaines et sociales que commence à apparaître et évolue les notions et pratiques des sciences humaines et sociales vers un nouveau concept qui est «*les humanités computing* » et puis «*les humanités numériques* ».

Ce phénomène récent, né au tournant du XXI siècle au carrefour des sciences humaines et sociales et l'informatique (Caraco, 2012), fait appel d'une part à des nouveaux rapports entre technologies numériques et sciences humaines et sociales, et relie d'autre part des professions distinctes et diverses disciplines à savoir celles liées à la documentation, les bibliothèques et les archives.

Cet article cherche donc à mettre l'accent sur l'articulation entre les humanités numériques et les archives. Il s'appuie sur un travail d'études de corpus autour des humanités numériques principalement en langue française et anglaise. D'autres textes sur les archives numériques ont été également examinés .Nous souhaitons ici examiner la place des archives à l'ère des humanités numériques, parlons-nous d'une révolution ou évolution dans le positionnement des archives et des archivistes ? Il a paru nécessaire de s'interroger sur l'impact des humanités numériques sur les archives, et savoir dans quelle mesure elles ont influencés sur les archives ,ses notions, ses principes, ses outils et fonctionnement , et surtout ses méthodes de conservation à long terme et en contre partie l'intérêt effectifs des archives à prendre part aux humanités numériques ?

Nous allons dans les sections suivantes donner quelques repères historiques sur les origines des humanités numériques, puis, étudier les impacts sur Les archives en s'intéressant à l'enjeu des archives et surtout la notion de conservation à long terme et la notion la continuité dans un contexte numérique caractérisé par l'abondance informationnelle.

Cet article constitue par conséquent un travail synthétique et préparatoire à une étude de recherche plus large sur les archives sous le prisme des humanités numériques.

I- les Humanités numériques :

1 –Les origines des humanités numériques :

Les origines des humanités numériques remontent aux années 40 (*Magué, 2014*). Dès lors, son apparition est marquée par une évolution passant par 3 grandes ères (*Burnard, 2012*).

1- "Literary and Linguistic Computing" LLC: 1940-1980

C'est l'époque de l'informatique littéraire et linguistique, époque marquée par l'utilisation des outils informatiques dans le domaine littéraire et linguistique. Ainsi que dans les recherches historiques d'où l'évolution de lexicométrie⁸³ et le textométrie⁸⁴ et l'utilisation des bases de données en histoire en introduisant les méthodes d'analyse quantitative.

Parmi les premiers corpus qui marquaient cette période Le *projet de "l'index Thomisticum"*⁸⁵ du père Roberto Busa, c'est une analyse lexicale complète des œuvres de Thomas d'Aquin qui consiste à taper sur des cartes perforées tout l'ouvrage de Thomas d'Aquin dans le but de créer un index automatisé dont le corpus peut être repéré et ses fréquences sont calculées.

2-les Humanités Computing HC:1980-1994

La transition de Literary and Linguistic Computing aux humanités computing est passée par l'émergence d'**un processus d'institutionnalisation** justifié par la création des centres, la mise en place des programmes et la fondation des associations en DH. Ainsi, cette période est marquée par l'apparition d'**une communauté scientifique structurée, cela** va de pair avec la création d'une liste de diffusion "Humanist"⁸⁶ en 1987 (*Burnard, 2012*) dans le but de fournir un forum de discussion de la propriété intellectuelle, scientifique, pédagogique, et d'échanger les informations et les questions sociales entre les participants. En

⁸³Lexicométrie : nf science linguistique qui étudie statistiquement l'emploi des mots c'est-à-dire combien de fois une personne prononce tel ou tel mot, dans quelle condition, etc(source : dictionnaire français lexicométrie)

⁸⁴Textométrie <http://textometrie.ens-lyon.fr>

⁸⁵*Corpus Thomisticum*, <http://www.corpusthomisticum.org>

⁸⁶ <http://dhhumanist.org/>

outre, à cet époque émergent **un besoin de standards** : tel que qui est la TEI⁸⁷ projet important de l'humanité computing.

3-Les Digital Humanities DH (1994....) :

Selon Lou Burnard l'expression « Digital Humanities » a marqué cette période qui se coïncide avec l'émergence du Web au milieu des années 1990(Burnard, 2012).Celà provoque des évolutions majeures aux humanités computing. C'est en 2004 que naquit le terme « Digital humanities » (Dacos, 2015) lors de la parution de l'ouvrage « A companion to Digital Humanities »⁸⁸. Le passage de computing aux digital (numérique) est justifié par l'influence du web qui a ajouté une nouvelle dimension à la sphère des humanités computing celle de la communication en ligne et le partage sur les médias sociaux (blogs, wikis, réseaux sociaux et outils de micro-blogging), le recours des chercheurs aux outils numériques est justifié par le besoin de diffusion des connaissances plutôt que pour l'analyse textuelles ou seulement l'archivage numérique. Dès lors tous les corpus doivent être disponibles en ligne à l'égard de : William Blake Archive⁸⁹ou des Van Gogh Letters⁹⁰.

2 –Définition des humanités numériques :

« Définir les *digital humanities* n'est pas une tâche aisée » affirme Benjamin Caraco (Caraco, 2012). Cette expression renvoie-t- elle au nouvelle discipline en SHS ?ou à une simple nouvelle appellation ou à une transformation radicale dans les méthodes de recherche en sciences humaines ou sociales ?

Le terme « humanités digitales » de l'anglais "*digital humanities*" est parfois aussi utilisé bien que l'adjectif « digital » désigne quelque chose « qui appartient aux doigts ». L'utilisation de l'expression anglophone *digital humanities* et son abréviation DH en français est couramment utilisé.

Selon la définition proposée par les participants de THATCamp, première conférence effectuée à Paris les 18et 19 mai 2010 : «les digital humanities concernent l'ensemble des Sciences

⁸⁷ TEI, Text Encoding Initiative, <http://www.tei-c.org/index.xml>

⁸⁸ Simens, Ray, John Unsworth, et Susan Schreibman.Companion to Digital Humanities <http://www.digitalhumanities.org/companion/>

⁸⁹<http://www.blakearchive.org/blake/>

⁹⁰<http://www.vangoghletters.org/vg>

humaines et sociales, des Arts et des Lettres. Les digital humanities ne font pas table rase du passé. Elles s'appuient, au contraire, sur l'ensemble des paradigmes, savoir-faire et connaissances propres à ces disciplines, tout en mobilisant les outils et les perspectives singulières du champ du numérique.(...) Les digital humanities désignent une transdiscipline, porteuse des méthodes, des dispositifs et des perspectives heuristiques liés au numérique dans le domaine des Sciences humaines et sociales » 91

3-Des actions dans le monde

Des nouvelles formes d'action se mettent en place à l'égard de THATCamp et Day of DH. Au contraire du modèle habituel du colloque universitaire s'impose le THAT Camp comme des « non-conférences », non conférences parce qu'ils se présentent sous forme des rencontres informelles et non hiérarchisées.

Les premiers THAT Camp sont nés à Washington (université George Mason) et puis le premier THAT Camp européen a été organisé à Paris en 2010 ⁹²puis 2012. Il a été suivi très rapidement par Cologne et Londres, et plusieurs pays partout dans le monde.

Le "Day of DH" est une initiative de Geoffrey Rockwell ; il s'agit d'une journée datant le 18 mars là où les participants se réunissent sur un blog toute la journée pour faire connaître leur idée, leur avis sur les humanités numériques⁹³, les données recueillies sont utilisées comme corpus de recherche.

II-Les archives à l'ère des humanités numériques :

1-Lien ancré entre les archives et les humanités numériques :

L'alliance entre les archives et les humanités en général n'est pas récent. Au début, les archives sont considérées comme sciences auxiliaires de l'histoire, en effet, elles sont considérées comme source historique "*dans la diversité de leurs formes, de leurs supports ou de leur typologie, ont servi de fondement à la connaissance de l'histoire*"(Mourier et Caya ,2011), et après,

⁹¹ THATCamp Paris 2010<http://tcp.hypotheses.org/318>

⁹² THATCamp Paris 2010<http://tcp.hypotheses.org/318>

⁹³<http://hypotheses.org/13656>

avec l'affranchissement de cette dépendance, les archives étaient ouvertes à d'autres disciplines encore faisant partie des sciences humaines et sociales comme la géographie, l'économie, la sociologie, etc.. qui utilisaient les archives pour approfondir leurs connaissances.

La relation dialectique avec les humanités numérique n'est pas récente, c'est le fruit de plusieurs changements que la discipline archivistique a subis, allant de l'utilisation des bases de données référentielles dans les services d'archives, vers l'utilisation des instruments de recherche en archivistiques jusqu'à l'avènement des archives numériques et l'évolution de statut de document et de profil des archivistes et l'évolution des théories et des fonctionnalités archivistiques dans un ère humaniste dite numérique.

2- des archives en mutation ? :

1- Les archives : évolution des notions ?

Les archives sont d'abord définies comme « *matériau à valeur historique* » (l'archivistique contemporaine). Elles sont définies par la suite par "Elsevier's Lexicon of Archives Terminology" comme étant « l'ensemble de documents reçus ou élaborés par une personne physique ou morale, publique ou privée et destinés par leur nature à être conservés par cette personne même ». De ces définitions, nous avons commencé à considérer les archives en tant que documents destinés à être conservés seulement. Ensuite la définition commence à s'élargir à « information organique et consignée » (Couture et Call, 1994) dans les vingt dernière années. Les notions des archives se sont développées avec la reconnaissance légale (exemple loi québécoise ou tunisienne) pour englober tous les documents produits ou reçus par une personne ou un organisme dans le cadre de ses activités.

Dans un environnement numérique, les archives ont progressivement signifié une "collection délibérée de substituts" (*purposeful collection of surrogates*). Elles ont passé d'une "collection d'objet matériels" qui peut être décrite dans les instruments de recherche traditionnels, et que ses matériaux sont rarement édités et annotés dans leur ensemble à un ensemble de "substituts numériques". Les significations changent avec le temps, et l'archivage dans un contexte numérique est venu de proposer l'alliance entre les fonctionnalités d'édition et d'archivage. (Kenneth, M.Price, 2009)

Il semble bien que ce soit le cas dans le domaine des humanités numériques là où la technologie de l'information a adapté "archives" pour désigner "des collections de données de sauvegarde"(collections of back up data)(Kenneth ,M.Price, 2009)

2- Mutation des procédures et des outils archivistiques :

Au début, les services d'archives utilisaient des outils de gestions traditionnels (registres, répertoires et fichiers papier). Avec l'informatisation des procédures, ces services ont changé leurs outils à l'égard des nouveaux instruments de recherche grâce à des progiciels d'archives qui permettent l'informatisation de la gestion documentaire.

Avec la normalisation archivistique, les outils et les instruments de recherche sont devenus plus homogènes et plus cohérents vue la structuration hiérarchique de documents qui est devenue encore très logique (Brigitte, 2006)

Avec le numérique et la tendance de développer les échanges de données, les recherches sont orientées vers la mise en place des normes indépendantes des plateformes logicielles, ce qui amène à l'adoption des normes internationales (ISO , AFNOR ...) et l'implication des formats d'export et d'échange (Dublin Core ...) et des métalangages SGML puis XML (eXtensibleMarkupLanguage) ainsi que des structures d'encodage tels que l'EAD (Encoded Archival Description) (Delpierre, 2012)

A l'ère des humanités numériques, les nouveaux outils et instruments de recherche archivistique facilitent le rôle des humanistes ; en effet, ces instrument offrent à l'historien deux modes d'accès complémentaires à la description des fonds d'archives : la navigation par le plan de classement et la recherche par mot dans le plein texte de la description, le premier permettant de pallier les faiblesses inévitables du second.. (Brigitte, 2006)

3- L'évolution du statut de document :

Le document dans sa définition traditionnelle est un sémaphore : signes inscrits, et imprimés sur un support (Muller, 2011) il est caractérisé par un double statut : l'objet matériel qui traduit une technologie et contenu culturel qui traduit une tradition.

A l'ère du numérique, les documents ont changé leur spécificité et nature passant d'un objet matériel à un objet immatériel. Ils ne sont *"plus perçus comme des objets matériels statiques, mais plutôt comme des concepts virtuels dynamiques. La création des documents ne repose plus sur des organisations hiérarchiques stables, mais elle se situe plutôt à l'intérieur de réseaux de fonctionnalités horizontaux et fluides"*.

Avec le numérique, le couplage signes et support est changé étant donné que le document numérique est à la fois une structure (langage informatique), forme (ensemble des données organisées) et signe (texte, image ou son) et médium (trace des relations sociales reconstruites par les dispositifs informatiques) (Muller, 2011)

Ainsi le centre de gravité de document est changé et déplacé du producteur au lecteur. *"Le lecteur qui, pour le lire, doit le recréer à partir des données et des métadonnées qui le constituent"*(Muller, 2011)

Avec les nouveaux types de documents qui ont été générés comme l'hypertexte, l'hypermédia, la structure de document est devenue plus complexe, étant donné que le contenu d'un seul document est lié grâce à des liens hypertexte à d'autres documents de forme et source et de contenu différents.

Dans le contexte de l'humanité numérique le concept de document s'élargit à la notion de "source", englobe tout ce qui peut être lu, *"Pour un historien, tout peut devenir document"* (Chabin, 2004); ainsi pas de frontière ou des limites dans la définition de document. Pierre-Édouard Portier énonce une notion totalitaire de document *"c'est un ensemble d'éléments interconnectés : des pages manuscrites numérisées, du texte transcrit et annoté, des collections de pages ordonnées, des zones d'images de manuscrits isolées, des relations nommées entre éléments quelconques,..."* (Portier, 2010)

4- Changement du cycle de vie du document : de la théorie des trois âges au "Record Continuum"

La théorie archivistique de trois âges a fortement marqué la pratique archivistique depuis sa formulation en 1950 par Theodore Schellenberg (Schellenberg, 1956). Le cycle de vie de

document⁹⁴ selon cette théorie passe par 3 périodes : active, semi active ou inactive. Le document est "actif s'il est indispensable au fonctionnement quotidien, donc utilisé fréquemment, et s'il doit en conséquence demeurer près de l'utilisateur, il est semi-actif s'il ne sert qu'occasionnellement, donc s'il est utilisé moins fréquemment et s'il peut être entreposé ailleurs que dans des espaces administratifs. Enfin, il est inactif s'il n'a plus d'utilité prévisible pour l'administration qui l'a créé ou reçu" (Couture, 1994)

Dans une démarche de gestion électronique de documents, de gestion des risques, émergent le "records management" dans les années 1950 et 1960 aux États-Unis⁹⁵, Résultats de nombreuses initiatives qui ont abouti à la publication de la norme ISO 15489 sur le records management en 2001, transposée en français en avril 2002 , fondé sur le concept de "cycle de vie" des documents, le records management permet de gérer le document "record" à l'âge courant et intermédiaire.

Sous l'impulsion de l'archivage numérique, les archives se trouvent appropriées à des nouvelles tendances et les rôles des archivistes doivent dépasser le rôle de gardien de la mémoire au rôle de contrôle et gestion depuis la création de document tout au long de son cycle de vie (Chebbi ,2012). Cela amène à la notion de "Records Continuum" (Canada, Australie), une nouvelle vision du document dans sa continuité dès sa création jusqu'à la conservation définitive, en exigeant la collaboration entre l'archiviste et l'informaticien aux différents aspects de travail archivistique.

5- Des nouvelles méthodes de traitements et fonctionnalités :

A l'ère de numérique les pratiques archivistiques ont connu une évolution remarquable, si traditionnellement la gestion des archives passe par des séries d'opération successive, tels que la collecte, l'évaluation et le tri ; avec le numérique ces trois séquences sont regroupées à l'origine même des documents (Delpierre, 2012).

⁹⁴Le Cycle de vie du document : création, diffusion et usage, gestion et maintenance, tri et élimination, préservation et conservation permanente

⁹⁵<http://www.enssib.fr/bibliotheque-numerique/documents/48452-le-records-management.pdf>

Avec le numérique se pose la question de nécessité de tri, si l'espace virtuelle de stockage est illimité quelle nécessité alors de faire la sélection ? Joël Poivre affirme positivement que dans un environnement numérique le tri même s'il est devenu assoupli il reste nécessaire non pour des raisons d'espace et de capacité de stockage physique limité mais pour des raisons de limitation de capacité humaine d'exploitation des archives (Poivre, 2004). Les critères de tri aussi ont changé " le tri n'est plus à faire entre ce qui est important et ce qui ne l'est pas – par rapport à quoi? Qui peut le dire pour l'avenir ?- mais entre ce qui est redondant et ce qui ne l'est pas" (Delpierre, 2012).

D'autre changement touchant les notions de versement, de cotation qui devient moins opérante, la mise en place d'un processus électronique de transfert rend les opérations de versement passant "au fil de l'eau" tel que mentionne Françoise Banat-Berger (Banat-Berger, 2009).

Dans ce processus électronique, la phase de versement est assurée par plusieurs normes telles que la norme PAIMS (Producer Archive Interface Methodology Abstract Standard, ISO 20652) (Banat-Berger, 2009). Cette norme " vise à définir pour l'archive et le producteur une méthodologie de travail commune et structurée en quatre phases" (Banat-Berger, 2009). Ces phases sont : la phase préliminaire qui consiste à faire la prise de contact et d'échange d'information entre le producteur et l'archive, la phase de définition formelle qui traite la négociation d'un protocole de versement présentant une identification précise et complète des objets numériques à livrer à l'archive par le producteur, la phase de transfert effectif des paquets d'information et la phase de validation des paquets versés.

Quant à la description archivistique, une nouvelle description s'impose à l'ère numérique. Si traditionnellement, la fonction "description" s'exerce sur les archives définitives, à l'heure des humanités numérique là où on construit et on décrit les corpus numérique ; la phase de description intervient dès la création de document numérique par le moyen des métadonnées. (Delpierre, 2012).

Les métadonnées telles que définies par Banat-Berger, Huc et Duploux "méta- provient du grec et signifie "après, au-delà de, avec", il exprime une idée de transcendance et les mots forgés à partir de "méta" appartiennent le plus souvent aux domaines de la philosophie et des sciences humaines. Ils signifient (ce) qui est au-delà de, ce qui dépasse et englobe la réalité désignée par

le 2^e élément. Ainsi métadonnées signifie "au –delà de données " qui dépasse les données " qui englobe les données" (Banat-Berger, 2009)

Dans cette optique numérique on peut conclure 4 types de métadonnées

- les métadonnées descriptives servent à organiser les connaissances contenues dans l'objet numérique par son identification, sa classification, et son hiérarchisation et qui permettent au gestionnaire "d'avoir une vision de la couverture d'une collection d'objets numériques" (Banat-Berger, 2009).
- les métadonnées techniques servent à identifier l'environnement technique des objets numériques et d'identifier et caractériser le format d'un objet numérique et cela pour des raisons prévisionnelles d'opérations de migration de format ou d'émulation.
- les métadonnées de structure permettent de connaître l'organisation de l'information de l'objet numérique par l'identification de son structure logique indiquant les liens avec les autres éléments ayant la même sens et physique indiquant la manière d'organisation et de stockage de l'information contenue dans l'objets numériques
- les métadonnées administratives servent à la gestion de la vie de l'objet numérique, "Elles regroupent les métadonnées d'identification, de contexte, de provenance, d'intégrité et de gestion de droits" (Banat-Berger, 2009).

La description des archives numériques doivent répondre à des pratiques standardisées comme : EAD (Encoded Archival Description) pour la description des archives et EAC-CPF pour la description des noms de personnes, de familles et de collectivités

Avec les humanités numériques, des nouvelles opérations ont vu le jour, ainsi les projets humanités numériques exigent certaines opérations techniques autour la création et l'organisation et la description des corpus numériques. Ces opérations consistent en la création de nouvelles collections par la numérisation qui forment un extrait des collections primaires, ensuite l'opération de transcription qui consiste à " *la création d'une chaîne formée de **la transcription** des caractères manuscrits, mais aussi l'annotation d'intervalles de cette chaîne, le découpage de zones dans l'image de la page manuscrite, l'association de ces zones à des intervalles du texte*"(Portier, 2010)Ensuite s'ajoute d'autres

opérations à l'égard du tissu des liens entre éléments quelconques du système.

6-la conservation à long terme des archives numériques

:

Conservier le contexte ?

À l'ère de numérique des questions qui se posent en ce qui concerne la conservation de contexte, selon Hans Booms, la conservation du contexte est primordiale dans son texte « Ordre social et constitution du patrimoine archivistique », affirme qu'il est primordial de conserver le contexte. D'autres auteurs insistent sur l'importance de conserver le contexte. Tel que, Luciana Duranti, dans « Archives as a Place » affirme que la conservation de contexte est un facteur qui garantit la fiabilité des archives et son authenticité est assurée par la conservation du contexte, par sa localisation et sa description, ainsi que par la permanence de leurs relations (Duranti 1995, 247)

Dans ce sens, les métadonnées participent dans la conservation du contexte, " *elle aborde le fait de conserver l'ordre original, le principe de respect du fonds, afin de conserver le contexte*"(Lecomte,2014)en effet si dans le cas des archives papier la conservation de contexte est assurée par la description du fonds, la localisation des documents. Dans le numérique, la conservation du contexte se fait par le biais des hyperliens, des relations entre source primaires et secondaires intégrées dans les métadonnées (Lecomte, 2014)

Conservier dans un modèle OAIS

La conservation des archives à long terme nécessite des modèles fonctionnels (OAI des formats et des différents langages XML (TEI, EAD, EAC)), et des métadonnées de préservation numériques (METS et PERMIS) ainsi qu'une mise en œuvre d'une stratégie de préservation numérique.

La Norme OAIS (Open Archival Information System) est un Système ouvert d'archivage d'information est un modèle conceptuel d'archivage de document numérique il permet la gestion, l'archivage, la pérennisation à long terme de documents numériques ; il est enregistré comme norme ISO sous la référence 14721:2003.

L'"OAIS," propose une vue globale cohérente de la question de l'archivage numérique" (Banat_Berger, 2009), par la détermination d'une liste complète de concept et des fonctions nécessaire au fonctionnement de l'archivage électronique.

Au plan fonctionnel il couvre toute les activités essentielles pour l'archivage numérique à savoir : les entrées (le mécanisme de versement), le stockage (la conservation des bits), la gestion des données (ajout, modification, suppression), l'administration et la planification de la pérennisation et enfin, l'accès (Banat-Berger, 2009). Dans un environnement OAIS la conservation de contenu des documents numériques nécessite un "conteneur conceptuel», en effet, à l'égard des documents papier conservés dans des boites en carton qui portent l'identification de document, les documents numériques nécessitent d'être conservés mais dans des "paquets d'information"⁹⁶.

Ces paquets comportent à la fois les données primaires (contenu d'information) et ses métadonnées qui concernent son identification, sa lisibilité, son enregistrement, ses modifications, son stockage, ses migrations, et les relations avec les autres objets numériques ainsi que les données utiles à la recherche (cote, intitulé, nom de producteur, conditions d'accès)⁹⁷

Ainsi dans le modèle OAIS la notion de délai de conservation se disparaît et la différence entre le court et le moyen et le long terme s'amenuise comme mentionne Banat-Berger "*Le modèle OAIS définit le long terme comme "une période suffisamment longue pour qu'il soit nécessaire de prendre en compte les changements technologiques, et notamment la gestion des nouveaux supports et formats de données ainsi que l'évolution de la communauté d'utilisateurs"* (Banat-Berger, 2009)

Préserver les documents numérique nécessite à la fois de penser à la question d'un modèle fonctionnel assurant la pérennisation des documents numériques, mais aussi de penser aux langages offrant des granites à long terme ; le XML (eXtensible Markup Language) à l'égard de SGML (Standard Generalized Markup Language) sont "*des langages de description à balises comportant un ensemble de règles, une syntaxe, des mots clés permettant de constituer des documents structurés*" (Banat-

⁹⁶http://www.patrimoine.vd.ch/fileadmin/groups/19/PDF/EXPO-2015_THEMES_TEXTES_01.pdf

⁹⁷ http://www.patrimoine.vd.ch/fileadmin/groups/19/PDF/EXPO-2015_THEMES_TEXTES_01.pdf

Berger, 2009). Le XML peut être aussi considéré comme format lisible par les machines et les humains. L'archivage des documents XML nécessite l'adoption de grammaires standardisées comme TEI (Text Encoding Initiative), (Banat-Berger, 2009)

La TEI selon Marin et Mounier est le plus important projet de l'époque de l'humanité computing, paru en 1987 (Martin et Mounier, 2015) puis développé pour devenir une norme d'encodage XML utilisée pour toute source textuelle numérisée qui vise à standardiser les pratique d'encodage et de structuration des textes (Martin et Mounier, 2015) avec la TEI se posent des question concernant les pratiques de structuration ,d'encodage de l'information et la diffusion et l'archivage des corpus.

Quant aux formats de préservation numérique il y'a le format PDF, formats ouverts et très répandus permettent de garantir la conservation à long terme et l'utilisation la plus simple de document et sa conservation de la présentation du document est aisée.

Des métadonnées pour la préservation numérique

Il est nécessaire de penser à la question de description de documents numériques afin de réaliser sa traçabilité , cela est assuré par les métadonnées de préservation numérique qui seront associées au données primaires dans les paquets d'information et qui présentent les identifiant et le contenu intellectuel de documents numériques, ses caractéristiques techniques, sa structure et son organisation , son contexte là où il est crée et l'historique des opération effectuées sur le document .

Parmi les standards pour l'expression des métadonnées de préservation il y'a : PREMIS (PREservation Metadata Implementation Strategies)

"PREMIS est un standard pour l'expression des métadonnées de préservation. Ce modèle concerne autant ceux qui conservent que ceux qui transmettent de l'information à conserver. Il définit principalement un dictionnaire de données utile pour décrire les métadonnées de préservation"⁹⁸

⁹⁸<http://www.archivesdefrance.culture.gouv.fr/gerer/archives-electroniques/standard/metadonnees/>

Ce standard est hébergé par la Bibliothèque du Congrès. Dans le même ordre de préservation pérenne il y'a le standard "METS" (Metadata Encoding and Transmission Standard) "METS est un standard définissant un format d'emballage, c'est-à-dire un format permettant d'organiser un ensemble d'informations liées de manière explicite. Ce standard définit un schéma XML qui est utilisé dans des systèmes d'archivage électronique pour donner une forme aux paquets d'informations, en particulier pour l'AIP (le paquet qui est conservé au sein du système)"⁹⁹

A l'égard de PERMIS il est hébergé par la Bibliothèque du Congrès.

Stratégie de préservation numérique

Dans le cadre du modèle OAIS, une stratégie de préservation numérique est basée sur des différentes méthodes utilisées plus concrètement pour la pérennisation de l'information à l'égard de la migration et l'émulation.

Justifiée par des raisons de dégradation des supports (moins fiables) et de l'obsolescence technologique (évolution constante) et de l'évolutions des besoins des utilisateurs ,La migration *"consiste en pratique à transformer un format de fichier en un autre ou à transférer des données d'un support d'enregistrement vers un autre sans modifier le contenu informationnel"*(Banat-Berger,2009)

La migration est un processus qui nécessite une planification passant par des différentes phases : l'étude d'opportunité (l'évaluation et la justification de migration), l'étude de faisabilité (l'évaluation et la mesure de risque), la spécification du processus (la mise en œuvre de processus de migration), la mise en œuvre (moyen de contrôle et l'évaluation des résultats).

L'émulation «est une technique qui consiste à élaborer un logiciel qui va être capable de simuler les services et les comportements d'un autre logiciel conçu pour une plate-forme différente. (Banat-Berger, 2009)

7-Les archives numériques : la mémoire et l'oubli

⁹⁹ ibid

Depuis longtemps, les archives dans leur contexte traditionnel sont considérées la mémoire de l'humanité. La mémoire humaine est basée sur l'oubli "*la mémoire à défaut d'oubli rend parfois le souvenir difficile*" (Pastinelli, 2009), oublier c'est effacer pour recevoir d'autres informations, c'est faire perdre pour reconstituer d'autres souvenirs, dès lors, se pose le problème de contradiction entre les archives qui permettent de reconstituer le passé par le tout conserver et l'oubli qui, sans lequel il devient impossible de constituer les souvenirs donnant sens au présent.

Selon Miladi Doueïhi, l'oubli est humain et c'est une force créative mais est ce que c'est le cas pour le numérique ? Il donne réponse «Mais la dimension humaine telle qu'elle est déterminée par le numérique n'a pas cette possibilité créative : tout est reporté dans un système rigide et autoréférentiel." (Vitali Rosati, 2011)

Avec le numérique se pose la question du dispositif numérique qui semble conserve la totalité du passé et englobe le continu du réel, or en réalité il fait la sélection.

Dans le numérique, il paraît que toutes les données qui transitent numériquement sont enregistrées et conservées, mais cette technique numérique n'échappe plus à l'oubli "*cette culture du tout numérique qui a réalisé le miracle de pouvoir virtuellement tout conserver aura in fine tout perdu ; par négligence et par désintérêt, bien plus que par incompétence*" (De Biasi ,2013)

8-Le nouveau profil des archivistes à l'ère des humanités numériques

A l'ère des humanités numériques peut-on parler des nouvelles compétences des archivistes à développer ?

D'abord les exigences de numérique dans la gestion archivistique a conduit les archivistes à orienter leur pratiques dans des nouveaux rôles professionnels, technologiques, sociales, ainsi, les missions des archivistes sont évoluées et son profil aussi.

A l'ère de numérique ,l'archiviste assume des nouvelles compétences à la fois archivistiques, historiques, et organisationnelles, à cet égard, Françoise Banat-Berger parle d'un rôle d'« *assistant à maîtrise d'ouvrage* »dans la mesure où il participe à l'avancement des projets de système d'information

de l'organisation et à la constitution des processus métiers, il assiste dans la réflexion sur l'environnement de document numérique, son traitement, sa descriptions , les métadonnées nécessaires à décrire, et sa conservation.

En outre, des compétences technologiques sont encore justifiés dans l'environnement numérique, en effet, l'archiviste conduit le projet informatique, met en place le système d'archivage numérique, il maîtrise les formats de données et de support ,et s'oriente vers les compétences technique et numériques (connaître le type d'encodage et métadonnées, les formats de conservation des données et logiciels dédiés) Muller mentionne qu'il devient " **un spécialiste de la gestion numérique de l'information**" (Müller, 2011)

Les projets des humanités numériques qui nécessite des pratiques de numérisation et des connaissances de constituer et de gérer les corpus numérique développent chez les archivistes les capacités et les compétences en matière d'infrastructures matérielles, d'architectures, de réseaux, ainsi que d'exploitation scientifique et humaines des archives numériques au profit des communiés scientifique. Ainsi ces pratiques des humanités numériques ont renforcé le rôle humain et sociale de l'archiviste qui a passé de son rôle de gardien de la mémoire et de l'entièreté de l'information à un agent actif qui participe à la mémoire sociale, et assure son exploitabilité et sa préservation à long terme.

Danaks, montre que le rôle de l'archiviste est changé d'une façon radicale "si dans la modernité il suffisait de garder tout dans un dépôt physique pour faire de «l'archivage», aujourd'hui dû aux technologies précisément et au développement du monde professionnel, il doit posséder les outils afin de s'assurer que l'information que les gens reçoivent soit intègre, impliquant sans aucun doute des mécanismes de contrôle" Aucune source spécifiée dans le document actif ».(Danaks, 2014)

Conclusion

Et si on recommence notre question « *Est-il possible qu'une chose soit à la fois numérique et humaniste ?* » dans les archives réside la réponse. Mémoire de l'humanisme, les archives qui résumant ses forces et ses faiblesses, ses souvenirs et ses oubliés, ses arts et ses littératures s'évoluent aux avancements

numériques pour jouer encore plus ce rôle justifiant à l'ère des humanités numérique une mémoire dite numérique.

Avec le numérique les archives ont subi des changements au niveau de leurs notions passant de « *matériau à valeur historique* » destinés à être seulement conservées, aux "substituts numériques" proposant l'alliance entre les fonctionnalités d'édition et d'archivage. A cet égard les documents ont changé leur statut, leur spécificité et nature passant d'un objet matériel statique à un objet immatériel virtuels dynamiques englobant des éléments interconnectés. Ce changement dans le statut des documents a engendré un changement de leur cycle de vie, passant par la théorie des trois âges au "Record continuum" (Canada, Australie) une nouvelle vision du document dans sa continuité dès sa création jusqu'à la conservation définitive.

Les pratiques archivistiques eux même ont connu une évolution remarquable, en effet dans une vision de "Record continuum" les séries d'opérations archivistique (collecte, évaluation, tri) passaient jadis successivement, sont regroupées à l'origine même des documents. La description archivistique, quant à elle intervient dès la création de document numérique par le moyen des métadonnées et des structures d'encodage (EAD, EAC.)

A l'ère des humanités numériques, les nouveaux outils et instruments de recherche archivistique facilitent le rôle des humanistes en l'offrant la navigation par le plan de classement et la recherche par mot dans le plein texte de la description. En contre partie les humanités numérique ajoutent au travail archivistique des nouvelles pratiques tels que la création de nouvelles collections, l'opération de transcription, l'annotation, le découpage de zones dans l'image de la page manuscrite, l'association de ces zones à des intervalles du texte". Ensuite s'ajoute d'autre opération à l'égard du tissu des liens entre éléments quelconques du système.

Sous l'impulsion des humanités numériques, on passe à un mode d'archivage dynamique qui force l'action des archivistes dès le début de la création de l'information, les archives se trouvent appropriées à procéder des nouvelles tendances et les rôles des archivistes doivent dépasser le rôle de gardien de la mémoire au rôle de contrôle et gestion.

Peut on parler encore d'un rôle de créativité de l'archiviste passant des archives aux données massives créés par des big

data, et des humanités numériques vers encore plus des humanités créatives ?

Bibliographie

Banat-Berger, Françoise, Laurent Duplouy et Claude Huc. 2009. L'archivage numérique long terme les débuts de la maturité, Paris: la Documentation française.

Brigitte, Mazon ;2006. « Les archives et la mutation informatique. », *Matériaux pour l'histoire de notre temps* 2/2006 (N° 82), p. 100-102 URL : www.cairn.info/revue-materiaux-pour-l-histoire-de-notre-temps-2006-2-page-100.htm).

Burnard, Lou., 2012. « Du literary and linguistic computing aux digital humanities : retour sur 40 ans de relations entre sciences humaines et informatique », in Mounier, Pierre, (dir.) *Read/Write Book 2 : Une introduction aux humanités numériques*. OpenEdition Press, 2012. <http://books.openedition.org/oep/242>. consultable en ligne :

Caraco, Benjamin. 2012. Les digital humanities et les bibliothèques. *Bulletin des bibliothèques de France* [en ligne], n° 2, [consulté le 07 mars 2016]. Disponible sur le Web : <<http://bbf.enssib.fr/consulter/bbf-2012-02-0069-002>>. ISSN 1292-8399

Chabin, Marie-Anne ,2004. Document trace et document source. La technologie numérique change-t-elle la notion de document? *Information - Interaction - Intelligence*, CEPADUES, 2004, 4 (1). <sic_00001020>

Chebbi, Aïda, 2012, *Archivage du Web organisationnel dans une perspective archivistique*, Thèse présentée à la Faculté des études supérieures en vue de l'obtention du grade de *Philosophiæ Doctor (Ph.D.) en sciences de l'information* Université de Montréal

Couture, Carol, Jacques ducharme et Jean-Yves Rousseau, 1988. L'archivistique a-t-elle trouvé son identité? .Argus, vol 17, n2, p.51-58

Clément, Jean, 2001. la littérature au risque du numérique

http://www.cairn.info/article.php?ID_ARTICLE=DN_051_0113

Djamel, Abdelkader Zighed, 2013. Les humanités Numériques en Sciences Humaines et Sociales.

https://www.ish-lyon.cnrs.fr/sites/www.ish-lyon.cnrs.fr/files/page/fichier/Zighed_2013_HumanitesNumeriquesSHS.pdf

Dacos Marin, Pierre Mounier,2015. Humanités numériques : État des lieux et positionnement de la recherche française dans le contexte international. [Rapport de recherche] Institut français. 2015, pp.9782354761097.

<https://hal.archives-ouvertes.fr/hal-01228945>

Danaks Patricia, 2014 Question interdisciplinaire : la philosophie et l'archives

https://www.academia.edu/10655438/Question_interdisciplinaire_la_philosophie_et_larchive._Automne_2014

De Biasi, Pierre-Marc, 2013 « Les archives de la création à l'âge du tout numérique », *Revue Sciences/Lettres* [En ligne], 1 | 2013, mis en ligne le 01 mai 2012, consulté le 08 mars 2016. URL: <http://rsl.revues.org/314>; DOI: 10.4000/rsl.314

Duranti, Luciana. 1995. Archives as a Place. In Paper Presented at a Half-Day Seminar in Sydney.

Delpierre, Nicolas ; Delpierre, Françoise ; Mirguet, Françoise. 2012 Les chantiers du numériques : Dématérialisation des archives et métiers de l'archiviste. Paris: Academia, 2012

Kenneth, M.Price, 2009 "Edition, project, Database, Archive, Thematic Research Collection:what'sin a Name? Digital Humanities Quarterly 3.3 (2009)

<http://www.digitalhumanities.org/dhq/vol/3/3/000053/000053.html>

Le Deuff, Olivier.2012 Humanisme numérique et littératies. Semen - Revue de sémio-

linguistique des textes et discours, Presses Universitaires de l'Université de Franche Comté (Pufc), 2012, pp.117-134.

http://archivesic.ccsd.cnrs.fr/sic_00776805v1/document

Le Deuff, Olivier 2015. Les humanités digitales précèdent-elle le numérique ? . Imad Saleh. H2PTM 15, Iste éditions, 2015, 978-1784051150.

<file:///C:/Users/fatma/Desktop/les%20humanités%20numériques/preprintledeuff.pdf>

Marcello Vitali-Rosati, Michaël Eberle-Sinatra. 2014 Pratiques de l'édition numérique. Collection « Parcours numériques »Collection « Libre Accès »224 pages • mars 2014

<http://parcoursnumeriques-pum.ca/histoire-des-humanites-numeriques>

Magué, jean Philippe,2014. Introduction aux humanités numériques, Présentation donnée aux séminaire HEG dédié aux Digital Humanities du 5 juin 2014.

<http://perso.ens-lyon.fr/jean-philippe.mague/other/talks/HEG2014/#/>

Mourier ,Jacques; Caya, Marcel ,2011.module 1:les archives pour quoi?

Müller, Bertrand 2011 .Archives et temps présent : considérations inactuelles ; Communication présentée au colloque Temps présent et contemporanéité, Paris, 24-26 mars 2011

Pastinelli, Madeleine ,2009 "La mémoire et l'oubli dans l'univers de l'archive totale.",EspacesTemps.net, Travaux, **19.02.2009**

<http://www.espacestemps.net/articles/archive-totale/>

Poivre, Joel, 2004. L'archivage électronique : principes, promesses et pièges

http://adeda78.free.fr/IMG/pdf/POIVRE_promesses.pdf

Portier, Pierre-Édouard, 2010. Construction des documents multistructurés dans le contexte des Humanités numériques

<file:///C:/Users/fatma/Desktop/les%20humanités%20numériques/a%20lire/Liris-4893.pdf>

Serge Noiret, 2013 Y a t-il une Histoire Numérique 2.0 ? sur Academia.edu :http://www.academia.edu/739198/Y_a_til_une_Histoire_Num%C3%A9rique_2.0 <http://cursus.edu/article/20611/comment-numerique-transforme-pratique-histoire/#.Vt7vWH3hCM8>

Schellenberb, Theodore.1956. Modern archives: Principles and Techniques; University of Chicago

Scholes R., Wulfman C., 2008, Humanities computing and digital humanities, South Atlantic Review, Vol.73, N°4, The changing university and the humanities, 50- 66

Vitali Rosati, Marcello, 2011. Une philosophie du numérique Lecture de Pour un humanisme numérique de Milad Doueïhi (Seuil, Paris 2011)

Bibliothèques numériques et crowdsourcing : analyses bibliométriques et text mining

Mathieu Andro (1,2), Imad Saleh (2), Samuel Szoniecky (2)

(1) DIST, Institut National de la Recherche Agronomique

mathieu.andro@versailles.inra.fr

(2) Paragraphe, Université Paris 8

Résumé

Dans le cadre d'une thèse traitant du crowdsourcing appliqué aux bibliothèques numériques, un corpus spécifique exhaustif et représentatif du sujet a été constitué. Cet article présente les analyses bibliométriques et text mining de ce corpus.

Mots clés : crowdsourcing, bibliothèques numériques, numérisation, bibliométrie, text mining, fouille de textes

A l'instar d'un nombre croissant d'organisations publiques et privées, les bibliothèques peuvent désormais bénéficier des contributions de leurs usagers grâce au *crowdsourcing*. Dans les bibliothèques numériques, le *crowdsourcing* reste actuellement principalement limité à la correction participative de l'OCR (Andro, 2015, 2). Quelques expérimentations de numérisation à la demande par financements participatifs (*crowdfunding*) ont également déjà été conduites (Andro, 2014).

S'il est généralement bénévole, le *crowdsourcing* en bibliothèque fait parfois aussi faire l'objet de récompenses ou de rémunérations via des plateformes. Il peut aussi être réalisé sous la forme de jeux. On parle dans ce cas de *gamification* (Andro, 2015, 1). Le travail peut aussi être réalisé

de manière involontaire sur le modèle des *reCAPTCHA*. On peut parler dans ce cas de *crowdsourcing* implicite.

Une thèse de recherche-action a été conduite autour de ce sujet au sein du Laboratoire Paragraphe. Au cours de ce travail, un corpus au sujet du *crowdsourcing* appliqué aux bibliothèques numériques a été construit. Dans le présent article, nous présentons les résultats d'analyses bibliométriques et *text mining* de ce corpus original et représentatif du sujet.

Méthodologie

Un corpus a été initialement et classiquement constitué à partir de recherches bibliographiques sur le Web of Science, ScienceDirect et Google Scholar. Les équations de recherche utilisées ont été améliorées de manière itérative au fur et à mesure que de nouvelles expressions ont été identifiées et nous ont permis de développer notre vocabulaire se rapportant à la numérisation, aux bibliothèques numériques, aux diverses formes de *crowdsourcing* évoquées en introduction, aux noms de projets, de plateformes, d'auteurs ou de revues spécialisées identifiés notamment grâce aux analyses bibliométriques.

Le vocabulaire utilisé a été géré à partir de l'outil de text mining Luxid WebStudio. C'est un outil de construction de vocabulaires à partir de corpus de textes. Le vocabulaire que nous avons construit comporte les termes précédemment évoqués sous les formes anglaises et françaises avec leurs synonymes. Afin d'en permettre l'éventuelle réutilisation, ce vocabulaire a été déposé en tant que donnée de la recherche au format XML SKOS sur la plateforme Zenodo (Andro, 2016, 1).

Le corpus a également été complété de manière plus classique à partir des bibliographies des publications que nous avons lues. Afin de ne pas avoir à répéter périodiquement nos équations de recherches sur nos sources au risque de faire des erreurs, de repérer plusieurs fois les mêmes publications ou d'en ignorer, nous avons également mis en place un dispositif de veille éditoriale avec l'outil DigiMind. Ce dispositif nous a permis d'enregistrer et de capitaliser les sources et le vocabulaire d'interrogation de ces sources et d'être alerté rapidement de toute nouvelle publication relative au sujet.

Les notices bibliographiques obtenues via ces divers moyens de collecte ont été importées sous EndNote, dédoublonnées grâce aux fonctionnalités de l'outil, nettoyées manuellement,

retravaillées avec Open Refine et analysées statistiquement avec Excel afin de produire les analyses bibliométriques.

Afin de produire les analyses, deux corpus ont été utilisés : un corpus de 219 références de la bibliographie de la thèse afin de produire les analyses bibliométriques déjà évoqués et un corpus de 1071 textes intégraux afin de produire les analyses text mining. Ce corpus contient les textes des 219 références de la bibliographie de la thèse mais également des publications sur le sujet qui ont été lues mais n'ont pas été retenues en bibliographie. Les analyses text mining ont été produites grâce aux outils Luxid. Cette technologie nous a permis de produire des analyses complémentaires à celles obtenues avec la bibliométrie à partir des champs structurés des bases bibliographiques.

Ces corpus constitués entre décembre 2013 et juin 2015 résultent d'une sélection par nos soins après lecture des publications. Ils diffèrent donc qualitativement des corpus généralement construits afin de réaliser des analyses bibliométriques et qui, n'ayant pas toujours fait l'objet d'une sélection manuelle, peuvent comporter du bruit ou du silence documentaires et qui ont pour effet de biaiser les résultats obtenus. Nous estimons que la méthodologie que nous venons de décrire nous a permis d'obtenir un corpus exhaustif mais aussi un corpus fiable et sans bruit, chaque publication ayant été lue. Dans ces conditions, nous estimons que les analyses produites ci-après sont bien représentatives du crowdsourcing en bibliothèque.

Résultats

Analyses bibliométriques

Quelle est la dynamique de ce sujet de recherche ?

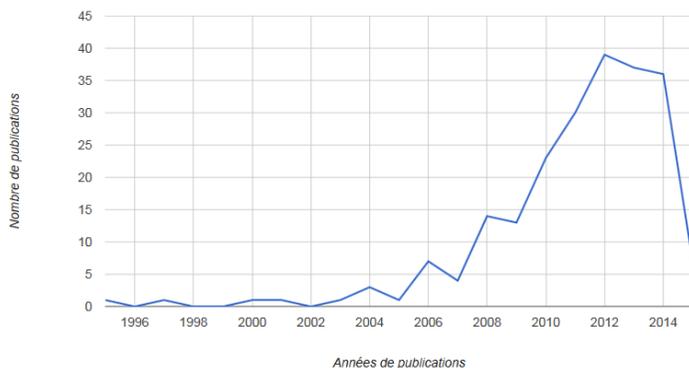


Figure 1. Nombre de publications par an dans la bibliographie de la thèse

Le sujet du crowdsourcing appliqué aux bibliothèques numériques est un sujet de recherche croissant, principalement à partir de 2006. Néanmoins, on observe une stagnation depuis 2013 qui coïncide probablement avec le passage d'une phase expérimentale à une phase plus industrielle.

Qui travaille sur ce sujet ?

Parmi les 361 auteurs du corpus, voici les plus spécialisés sur le sujet :

Holley R. 9 occurrences	Deterding, S 3
McKinley, D. 7	Dijkshoorn, C 3
Von Ahn, L 7	Gstrein, S 3
Aroyo, L 4	Ipeirotis, P. G 3
Smith-Yoshimura, K 4	Mühlberger, G 3
Alam, S. L 3	Tonra, J 3
Blum, M. 3	Wallace, V. 3
Campbell, J. 3	

Parmi les 149 institutions d'auteurs de notre bibliographie de thèse, voici les plus représentées :

Australie - National Library of Australia – 11 occurrences
Nouvelle Zélande - Victoria University of Wellington - 9
USA - Carnegie Mellon University - 8
UK - University College of London - 6
Danemark - University of Copenhagen - 5
Autriche - Innsbruck University - 4
Pays Bas - University Amsterdam - 4
USA - OCLC - 4
USA - University of Illinois – 4
USA - University of Iowa - 4

Parmi les 31 pays de ces auteurs, on obtient la répartition suivante sur une carte :

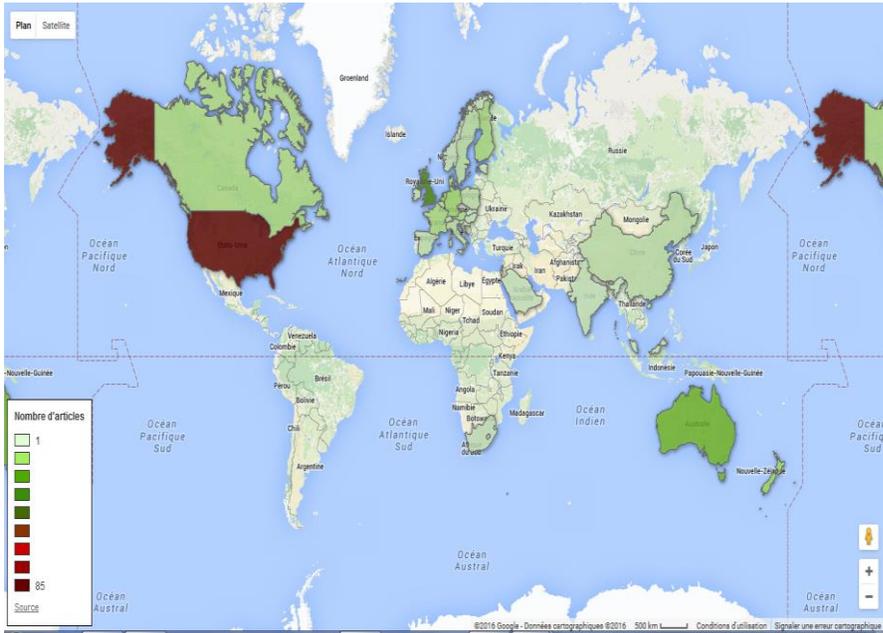


Figure 2. Poids des pays des auteurs dans la bibliographie de la thèse

On observe une nette domination des pays anglo-saxons, Etats-Unis puis Royaume Uni en tête. On observe aussi l'absence des pays en voie de développement et des pays du Sud. Concernant la France, nous n'avons considéré que les textes en langue internationale afin de ne pas biaiser les résultats.

Analyses text mining

Quels types de *crowdsourcing* sont évoqués ?

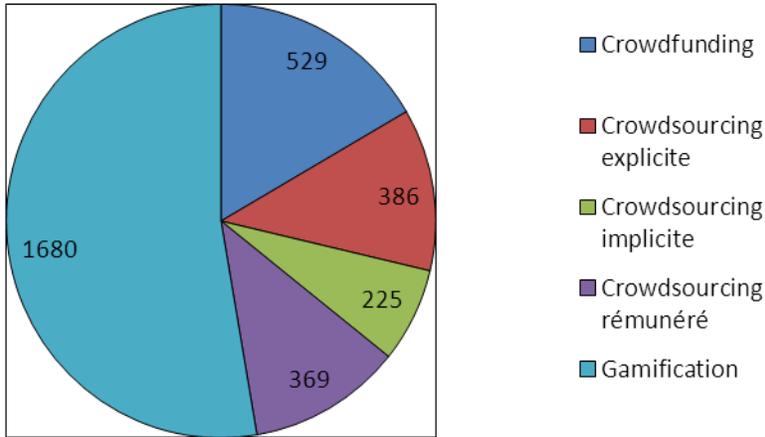


Figure 3. Nombre d'occurrences par type de crowdsourcing dans le corpus text mining

La gamification est un domaine très représenté dans la littérature scientifique. Il semble intéresser la recherche scientifique bien que le nombre de projets en bibliothèque reste encore assez faible par rapport aux projets classiques de *crowdsourcing*. Le *crowdsourcing* explicite reste, en effet, très majoritaire mais semble moins intéresser le monde académique du fait de son caractère désormais moins innovant. La même observation peut être faite aussi pour le *crowdfunding*, le *crowdsourcing* rémunéré ou le *crowdsourcing* implicite qui sont encore peu expérimentés en bibliothèques.

Quels types de motivations sont évoqués ?

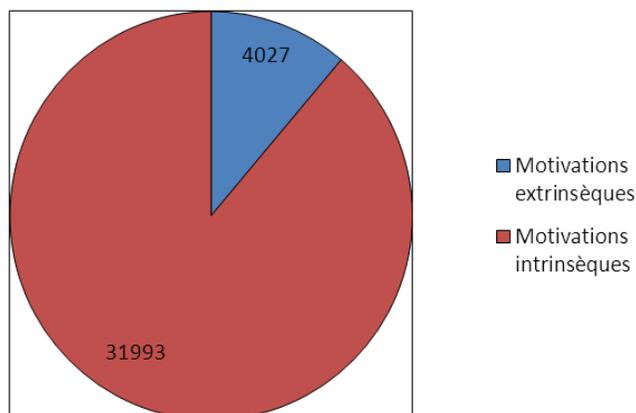


Figure 4. Nombre d'occurrences relatives pour les motivations extrinsèques et intrinsèques dans le corpus text mining

Nous définissons les motivations intrinsèques dans le domaine du *crowdsourcing*, comme des motivations qui poussent les internautes à contribuer pour le plaisir et le seul intérêt du travail proposé de manière désintéressée. Au contraire, les motivations extrinsèques les poussent à contribuer afin d'obtenir un bénéfice extérieur au travail lui-même comme la reconnaissance, les récompenses symboliques, matérielles ou financières.

Ce dernier type de motivation semble moins étudié. Le crowdsourcing rémunéré, en particulier, a encore très peu été expérimenté et étudié dans les bibliothèques.

Conclusion

Les analyses que nous avons produites portent sur un corpus qui nous semble le plus exhaustif et fiable chaque publication ayant été consultée et sélectionnée. On observe une croissance de la littérature sur le crowdsourcing appliqué aux bibliothèques numériques, en particulier dans les pays anglo-saxons. Le domaine de la gamification intéresse, en particulier, beaucoup les communautés scientifiques. Celui du crowdsourcing rémunéré nous semble encore avoir fait l'objet d'assez peu

d'expérimentations. Il serait donc judicieux d'en envisager l'expérimentation prochaine.

Bibliographie

Andro, M., Rivière, P., Dupuy-Olivier, A., Gropallo, F., Maingreud, D. (2014). Numalire, une expérimentation de numérisation à la demande du patrimoine conservé par les bibliothèques sous la forme de financements participatifs (crowdfunding). Bulletin des Bibliothèques de France, contribution du 2 octobre 2014, 9 p.

Andro, M., Saleh, I. (2015, 1). Bibliothèques numériques et gamification : panorama et état de l'art. I2D - Information, données & documents. 52(4), p. 70-79.

Andro, M., Saleh, I. (2015, 2). La correction participative de l'OCR par crowdsourcing au profit des bibliothèques numériques. Bulletin des Bibliothèques de France, Contribution du 16 juin 2015.

Andro, M. (2016, 1). Crowdsourcing thesaurus. <http://dx.doi.org/10.5281/zenodo.59507>

Archives et transition démocratique à l'ère de des humanités numériques : la situation en Tunisie post-révolutionnaire

Yousra Seghir, Khaled Habchi

Unité de recherche BNP, ISD

yousra.seghir@gmail.com, Khaledhabchi@hotmail.com

Campus universitaire de la Manouba, 2010 Manouba-Tunisie

Introduction

Dans les pays ayant connu des régimes autocratiques et vécu de longues périodes sous la dictature comme la Tunisie, de graves violations des droits humains et du droit humanitaire international sont commises tous les jours. La corruption et la malversation y trouve un terrain favorable pour se propager. Les documents d'archives sont les témoins de ces violations et constituent une valeur de preuve et permettent d'établir les responsabilités et réparer les victimes. Dans cet ordre d'idée, les fonds d'archives des institutions répressives ainsi que les archives des ONG de la société civile œuvrant dans le domaine de la défense des droits de l'Homme revêtent une importance capitale, car ils représentent une source incontournable dans le processus de la justice transitionnelle. La récupération de ces archives, leur traitement et leur exploitation sont la clé de voute de tout ce processus. La mise en place et la tenue d'un système de gestion des documents d'archives sur les violations en Tunisie, constituera un soutien inestimable à la démocratie naissante et empêchera que des violations ne soient commises et ne se répètent.

1. Problématique

La Tunisie a connu depuis plus d'un demi-siècle un régime totalitaire sous lequel des violations des droits de l'Homme de toutes sortes ont été perpétrées. La corruption est devenue monnaie courante et a même pris la forme de corruption d'état. Le soulèvement tunisien du 17 décembre/14 janvier, qui a eu

pour conséquence la chute du pouvoir en place, a permis de rompre avec la dictature et de jeter les jalons d'une jeune démocratie. La transition démocratique que connaît le pays suite à la révolution du 14 janvier 2016, entamée sous le signe de la réconciliation nationale, a permis de reconnaître le rôle et l'importance de la conservation des documents d'archives relatifs à ces violations et à ces abus, car ils présentent *«bien des preuves de ce qui s'est passé et construit [...] le fondement de ce qui doit se faire, la preuve aussi de la situation réelle»*¹⁰⁰.

L'importance des archives a été vérifiée à l'occasion des travaux d'enquête entrepris par les commissions de vérité créées dans le cadre de la justice transitionnelle. Des milliers de plaintes, ont été déposées à l'Instance de Dignité et de Vérité (IVD). L'IVD a éprouvé beaucoup de difficultés à rassembler les fonds d'archives indispensables à son travail d'investigation, car l'essentiel de ces documents est hébergé dans les locaux de la présidence de la république et du ministère de l'intérieur. D'autres fonds d'archives ont été vandalisés, d'autres ont tout simplement disparu. En effet, durant les premiers jours de la révolution, conscients de leur importance, les symboles de l'ancien régime ont tout simplement détruit les documents d'archives pouvant les incriminer. Les dossiers retrouvés ne répondent pas tout à fait aux normes de conservation et d'organisation des documents d'archives, bien que la Tunisie dispose d'un système national de gestion des documents et des archives performant.

Pour jouer pleinement leur rôle dans la transition démocratique et au-delà, les archives sensibles liés aux droits humains, aux libertés et à la corruption devraient faire l'objet d'une conservation adéquate et d'une organisation rationnelle afin d'en assurer la pérennité.

La présente étude essaiera de répondre aux questions de recherche suivantes :

- Comment préserver les archives et le patrimoine documentaire relatifs aux questions de violations des droits de l'Homme et assurer leur pérennité ?
- En tant que mécanisme essentiel de la justice transitionnelle, quel est le rôle et les fonctions de

¹⁰⁰ Bazin. Paule René «Rôle et limites des archives dans la construction de l'histoire.». *Revue Quart Monde*, N°199 - "Forger la mémoire d'un avenir commun", Année 2006
Revue Quart Monde

l'activité archivistique au sein de ce processus et plus tard dans le contexte d'une gouvernance de connaissance et administrative transparente ?

- Qu'en est-il de la communicabilité et la consultation de ces archives ?

2. De la transition démocratique

La notion de transition démocratique est un concept né avec la vague de révolutions démocratiques qu'a connu l'Europe du sud (Espagne, Portugal, Grèce) dans les années 1970 et ensuite, l'Amérique latine et l'Europe centrale et orientale dans les années 80 et 90. Elle désigne une « *modalité de changement politique* »¹⁰¹ marquant le passage progressif d'un régime autoritaire à un régime démocratique. Elle est nécessaire dans la mesure où elle « *entraîne l'abandon des anciennes règles du jeu politique et suscite l'apparition de nouveaux acteurs politiques et de nouvelles configurations stratégiques* »¹⁰². En ce qui concerne la Tunisie, cette phase est aussi importante que nécessaire. Le régime politique révolu est hautement autocratique et répressif ce qui a causé un volume très important d'abus, de pratiques dégradantes et de violations des droits humains. La population est dépourvue de traditions et de culture démographiques. Cette phase de transition se traçait comme objectif de « *reconstruire le pays et instaurer un climat de paix et de réconciliation* »¹⁰³. Elle repose sur le processus de la justice transitionnelle dans une quête de faire table rase du passé, établir les responsabilités et garantir la réparation des victimes de violation.

Sous la pression d'une société civile forte et dynamique, des instances de vérité furent créées. Leur rôle est de « *mettre en œuvre un mécanisme de justice transitionnelle, c'est-à-dire dans un cadre légal, inspiré par l'évolution de la théorie de la justice transitionnelle et sa pratique dans d'autres pays, et qui s'inscrit dans la nouvelle Constitution (adoptée en*

¹⁰¹ Akindes, Francis.- Rapport introductif n° 3 : Les transitions démocratiques à l'épreuve des faits : réflexions à partir de expériences des pays d'Afrique noire francophone».- in : Actes du Symposium international sur le bilan des pratiques de la démocratie, des droits et des libertés dans l'espace francophone, Bamako (Mali), 1er - 3 novembre 2000. P. 609

¹⁰² Linz, Juan J. Democratic transitions and democratic consolidation, mimeo, July 1991, p.2.

¹⁰³ Les Archives et le droit de savoir : actes de la journée de réflexion organisée à l'occasion de la journée nationale des archives par l'UNESCO et le Archives nationale de Tunisie, 2013, P. 6

janvier 2014)»¹⁰⁴. La principale instance de vérité mise en place est l'Instance de Vérité et de Dignité »¹⁰⁵. Au 31 mai 2016, elle a reçu plus de 24000 plaintes de la part de personnes victimes de violations de droits humains, sociaux et économiques. L'IVD est la principale utilisatrice des archives sur les violations. Selon son statut juridique, elle dispose de tous les droits pour accéder aux archives des instances répressives couvrant la période de juillet 1955 à décembre 2013.

3. Archives et humanités numériques

Le domaine des archives a été longtemps considéré comme une branche de l'histoire. Les documents d'archives servent à éclairer les chercheurs dans cette discipline et dans les sciences humaines et sociales (SHS) en général. Le développement vertigineux des TIC a permis aux archives de franchir les « *frontières numériques* »¹⁰⁶, conforter et consolider ainsi leur statut comme documents incontournables en matière de recherche scientifique en SHS. On parle désormais de corpus de document d'archives et de portails d'archives à l'instar du corpus de la recherche en sciences humaines et sociales, le portail européen des archives¹⁰⁷ et le portail international archivistique francophone (PIAF)¹⁰⁸. Ces outils permettent outre une conservation pérenne et sécurisée, un accès local ou distant aux ressources consignées.

En Tunisie, l'information numérique connaît un développement vertigineux depuis les années 1990 du siècle dernier. Tant au niveau de la recherche qu'au niveau de l'administration, les corpus documentaires se multiplient. L'administration électronique est pionnière dans la région Afrique du nord/

¹⁰⁴ INTERNATIONAL CRISIS GROUP.- Rapport Moyen orient/Afrique du nord, 23 mai 2016

¹⁰⁵ INTERNATIONAL CRISIS GROUP.- Tunisie : justice transitionnelle et lutte contre la corruption : rapport Moyen orient/ Afrique du nord, n° 168, mai 2006.- p.i

¹⁰⁶ Saleh, I. ; Bouhai, N. ; Achour, H.- *Les frontières numériques*. Paris : L'Harmattan, 2014.- 252 p.

¹⁰⁷ <https://www.archivesportaleurope.net/fr/home>

¹⁰⁸ www.piaf-archives.org/

moyen orient et en Afrique. En matière d'information numérique judiciaire, il y a lieu de citer les projets de numérisation des dossiers de justice dans les tribunaux, entrepris dans le cadre du projet de modernisation du système judiciaire tunisien (AMSJ) en 2008, financé par l'Union européenne (UE). Un autre projet de même nature est relancé en 2015 avec l'aide de l'UE.

4. Les archives sur les violations, archives de la dictature : des archives « sensibles »

Les archives sur les violations sont considérées comme des archives sensibles. A l'origine, ce sont des documents importants ou « vitaux » contenant des informations confidentielles ou secrètes. Tous les secteurs d'activité ont leurs lots d'archives sensibles et secrètes. En France, « *chaque fonds d'archives possède [...] ses documents réservés, éventuellement soumis des formes de dérogation particulière* »¹⁰⁹. Ces archives sont évaluées et classées par les services créateurs selon le « *sens que l'on donne au mot sensible* »¹¹⁰.

En Tunisie, les archives de la police politique et autres institutions répressives et de propagande politique dont les archives sur les violations sont considérées comme des archives « *ultra-sensibles* » et sont jalousement gardées dans les locaux des institutions créatrices. F. Hachad et al. (2015) prête à ce type d'archives le nom d' « *archives de la dictature* » et le définit en tant que « *documents ayant servi la stratégie consistant à installer, renforcer et perpétuer la dictature. Par exemple, les archives sécuritaires, qui en démocratie, sont supposés garantir la sécurité humaine, prennent une autre ampleur sous la dictature et servent d'instrument de chantage ou de répression : écoutes téléphoniques, espionnage* »¹¹¹. Pendant le soulèvement populaire du 14 novembre 2011, une partie de ces archives furent détruites, notamment les archives de la présidence et du Rassemblement Constitutionnel Démocratique (RCD) et de 13 tribunaux, de postes de police. Mais « *il n'y a pas eu de pertes*

¹⁰⁹ DULLIN, Sabine.- Le secret et ses usages : les archives « sensibles » en URSS. P.189. -<http://books.openedition.org/éditions-CNRS/1520?lang=fr>

¹¹⁰ Ibid.- p. 189

¹¹¹ Hachad, Farah ; Ferchichi, Wahid, dir. Révolution tunisienne et défis sécuritaire. T1 : Les Archives de la dictature entre justice transitionnelle et sécurité

notables au niveau des archives puisque tout a été regroupé dans les ministères et centralisé»¹¹².

Quelques ONG des droits de l'Homme et de la société civile détiennent de leur part des archives sur les violations issues des plaintes des victimes. Quelques une d'entre-elles disposent même d'archives numériques structurées sous forme de bases de données sur les violations et utilisent des formats standards d'enregistrement et d'échange de données.

La question des archives sensibles sur les violations ou archives de la dictature a émergé durant la période transition démocratique qui a suivi la révolution tunisienne de 2011 et a suscité beaucoup de débat. Ces archives ont généré nombre d'interrogations en raison de leur extrême importance dans le processus de la justice transitionnelle. Les instances *ad hoc* d'investigation et les commissions de vérité ont fait appel aux archives pour traiter les dossiers des plaignants victimes de violations. Le débat s'est articulé sur le droit d'accès et de consultation et sur la valeur de ces archives en tant que pièces à conviction. Les informations qui y sont contenues représentent des indices qui doivent être recoupés et validés.

5. Droit d'accès à l'information et au savoir consignés dans les archives sur les violations

Le droit d'accès à l'information, aux documents administratifs et au savoir est un droit reconnu dans la plupart des pays démocratiques. Il constitue une partie intégrante des droits de l'Homme. Dans les pays autocratiques, ces droits sont bafoués voire inexistantes.

S'agissant de la Tunisie de l'après révolution, un arsenal juridique garantissant au citoyen ce droit, désormais constitutionnel, a été mis en place dès 2011, le décret-loi 41 est venu régir le droit d'accès aux documents administratifs. En mars 2016, le projet de loi organique sur l'accès à l'information fut adopté par l'Assemblée des représentants du peuple (ARP), en attendant d'entamer l'élaboration des textes d'application y afférents. Les archives sensibles sont régies par loi n°1988-95 datée du 8 Août 1988 qui repose sur le principe de délais de communicabilité. Ce principe est la pierre angulaire du système d'accès aux archives. Son principe vise «à ce que les documents publics ne pourraient être communicables qu'à l'expiration

¹¹² Ibid

d'un certain délai qui diffère d'un pays à un autre et aussi selon les catégories des documents »¹¹³. Pourtant, les principes relatifs à l'accès aux archives établis par le Conseil International des Archives (CIA) reconnaissent le droit aux victimes de violations d'accéder aux archives pour constituer des preuves à l'encontre de responsables de ces violations. Elle recommande aux institutions conservant ce type d'archives de faire en sorte que «les victimes de crimes graves ressortissant du droit international aient accès aux archives qui fournissent les éléments de preuve nécessaires pour faire valoir les droits de l'Homme et documenter leur violation, même si ces documents ne sont pas communicables au grand public»¹¹⁴.

6. Droit de savoir, préservation de la mémoire

La problématique de droit d'accès aux archives sensibles en Tunisie ne concerne pas seulement les victimes des violations, elle concerne également le reste de la population. De ce point de vue, elle traverse le territoire du « *personnel* » pour franchir l'espace patrimonial. Il s'agit d'un droit de savoir et d'une prise de connaissance de ce qui s'est réellement passé durant l'ère de la dictature. Le rapport 2005 de la Commission des droits de l'Homme des Nations Unies souligne que «*la connaissance par un peuple de l'histoire de son oppression appartient à son patrimoine et, comme telle, doit être préservée par des mesures appropriées au nom du devoir incombant à l'État de conserver les archives et les autres éléments de preuves se rapportant aux violations des droits de l'Homme et du droit humanitaire et de contribuer à faire connaître ces violations. Ces mesures ont pour but de préserver de l'oubli la mémoire collective, notamment pour se prémunir contre le développement de thèses révisionnistes et négationnistes*».

7. Méthodologie de recherche

Pour répondre aux questionnements sur les archives spécifiques et la transition démocratique, notre étude se penchera sur l'examen de la situation des archives relatives aux violations des droits de l'Homme en Tunisie dans la période suivant la

¹¹³ Saidi Mohamed.- Communication des archives définitives.- www.archivesdefrance.culture.gouv.fr/static/2986

¹¹⁴ Conseil International des Archives : *Principes relatifs à l'accès aux archives établis par le Conseil International des Archives*, adoptés par l'AGM le 24 mai 2012.- p. 11

révolution du 14 janvier 2011. Dans un premier temps, Un état des lieux des fonds d'archives portant sur ces questions sera établi à partir d'entretiens semi directifs réalisés pendant la période allant du 1^{er} au 10 septembre 2016 et auprès des instances concernées qui sont au nombre de six institutions et instances des droits de l'Homme et des libertés représentant les organismes publics, les commissions ad hoc et les ONG de la société civile. Ensuite, elle proposera les solutions nécessaires aux carences constatées. Dans un souci de préservation des archives et du patrimoine documentaire, un système de gestion de documents d'archives évolutifs vers de systèmes de gestion des connaissances sera proposé. Les aspects techniques tels que la conservation et la gestion et d'autres à connotation politique tels que l'accès, la propriété et l'hébergement seront tour à tour étudiés et discutés.

8. Analyse des résultats de l'enquête

La période post-révolutionnaire en Tunisie a vu se développer l'idée du dévoilement des archives afin de faire surgir aux citoyens la vérité. Delà un intérêt particulier aux archives relatives aux violations des droits de l'Homme s'impose étant donné que ces archives « deviennent dans le nouveau régime politique, avec l'arrivée des libertés et à la lumière de la référence commune qu'offre la déclaration universelle des droits de l'Homme, un instrument social irremplaçable pour établir de nouvelles relations sociales »¹¹⁵. A l'ère de la transition démocratique et de la justice transitionnelle, les archives renforcent l'exercice des droits de l'Homme et consolident la démocratie. Il s'agit de réfléchir, dans cet article, sur le tournant des mécanismes de gestion des archives à l'ère des humanités numériques pour accroître la production, le partage, la diffusion et la valorisation du savoir et du savoir-faire contenus dans des documents d'archives. Afin d'analyser la situation des archives sur les violations des droits de l'Homme en Tunisie, nous avons interviewé six responsables des différentes structures mentionnées dans le tableau suivant :

Structures	Nature de la structure	Date de création	Personnes interviewées
Archives Nationales	Institution Nationale	1988	Directeur Général

¹¹⁵ Quintana, A. G., *Politiques archivistiques pour la défense des droits de l'Homme*, Paris, Conseil international des archives, 2009

de Tunisie (ANT)			
Instance de Vérité et de Dignité (IVD)	Constitutionnelle	2014	Membre élu spécialité Archivistique
Comité Supérieur des Droits de l'Homme et des Libertés Fondamentales (CSDHLF)	Institution nationale	2016	Conseiller auprès du président du comité
Commission nationale d'investigation sur les dépassements et les violations (CNIDV)	<i>Ad hoc</i>	17 décembre 2010- 14 janvier 2011	Membre de la commission
Ligue Tunisienne des Droits de l'Homme (LTDH)	Société civile	1977	Membre du bureau de la Ligue
Association Tunisienne des Femmes Démocrates (ATFD)	Société civile	1989	Membre du bureau de l'association

Tableau 1 : Structures interviewées

Nous avons choisi différentes structures qui détiennent les archives sur les violations des droits de l'Homme. Les archives nationales de Tunisie ont pu récupérer les archives des institutions répressives telles que : le RCD (Rassemblement Constitutionnel Démocratique), ATCE (Agence Tunisienne de Communication avec l'Extérieur), Le Ministère de la communication, la Chambre des conseillers (les archives de la commission des affaires politiques, des droits de l'Homme et des relations extérieures), la CNIDV (Commission Nationale d'Investigation sur les Dépassements et les Violations) et ce, pour préserver ce patrimoine et le protéger contre toute destruction. Les autres structures, que ce soit instance constitutionnelle ou nationale ou de la société civile, étant donné qu'elles détiennent des archives importantes qui révèlent la vérité des violations. Il est à noter que la loi organique relative à

l'instauration de la justice transitionnelle et à son organisation, dans son article 4, stipule que :

«Le dévoilement de la vérité englobe l'ensemble des moyens, procédures et enquêtes mis en œuvre pour démanteler le système despotique et cela, à travers la délimitation de toutes les atteintes, leur identification, la connaissance de leurs motifs, de leurs circonstances et les conditions dans lesquelles elles se sont produites, de leurs auteurs, des résultats qui en découlent, de leur emplacement, des auteurs de ces actes et de ceux qui en sont responsables».

Delà, la réception et la constitution des documents d'archives par ces structures ne peuvent que renforcer l'atteinte de leurs objectifs. En effet, selon les lois organiques qui régissent leur fonctionnement, ces structures cherchent à :

- Instaurer, mettre en œuvre les mécanismes de la justice transitionnelle et collecter les données, repérer, recenser, confirmer et archiver les violations en vue de constituer une base de données et d'élaborer un registre unifié des victimes de violations (IVD) ;
- Recevoir, faire le suivi des plaintes et traiter les informations relatives aux violations des DH pour constituer par la suite des archives (CSDHLF) ;
- Enquêter sur les exactions et violations commises au cours de la période allant du 17 décembre 2010 (CNIDV) ;
- Défense des droits de l'Homme, accompagnement des victimes de violences, accompagnement judiciaire aux victimes pour les aider à couvrir les frais des avocats et formation à l'application des règles des droits de l'Homme (LTDH) ;
- Faire des études et des analyses sur les violations commises sur les femmes et constituer une mémoire sur les violations sur les femmes (ATFD).

En effet, ces structures produisent une masse énorme de fonds archivistiques résultant des pratiques de la violence sur des personnes physiques et morales et qui témoignent aussi de la responsabilité de ceux qui ont commis les agressions. Ces archives sont considérées comme une source de savoir qui nous renseigne sur le passé et nous permettent de s'ouvrir sur des traces de la vérité historique.

L'importance de ces archives nous a permis de s'interroger sur les modalités de leur constitution, de leur traitement, de leur

préservation classique et digitale et de leurs modes de communication. En plus des entretiens avec les responsables de ces structures, nous avons fait le tour des locaux dans lesquels ces archives sont conservées et nous avons pu observer l'état de la tenue de ces documents essentiels.

9. La localisation des fonds d'archives sur les violations des droits de l'Homme

Pour garantir la disponibilité des fonds d'archives et la préservation de leur trace en cas de disparition, les instances devraient disposer de locaux dédiés à leur conservation. Malheureusement, ce que nous avons pu remarquer à travers le questionnement de la situation des archives que, mise à part les archives des instances répressives, qui ont été récupérées par les archives nationales, la majorité des structures possédant des archives importantes ne dispose pas de locaux de conservation normalisés. Le tableau suivant illustre bien cette situation :

Instances	Les fonds d'archives	locaux
ANT	- Les procès politiques depuis les années 60. - Fonds du RCD. - Fonds d'archives CNIDV	Magasins des Archives nation
IVD	- Collecte des archives publiques et privées (copies) auprès des structures suivantes : la présidence, le ministère de l'intérieur, le ministère de la justice. - Les enregistrements audiovisuels des témoignages	Locaux selon les normes l'accès est avec empreinte caméras de surveillance e conservations selon les norme
CSDHLF	Les archives sur les plaintes, les rapports et les documents qui concernent le suivi des plaintes (courriers ayant une relation avec les violations)	Dans les locaux du CSDHLF
CNIDV	Les archives ont été transférées aux archives nationales	Magasins des Archives nation
LTDH	Les archives sur les violations des droits de l'Homme : des plaintes des victimes	Un noyau des archives se trou la ligue Des archives stockées dans le membres militants jusqu'à no l'achat d'un local propre à la
ATFD	Les archives sur les femmes violentées	Les archives sont stockées au

Tableau 2 : Localisation des fonds d'archives sur les violations des droits de l'Homme

En effet, les fonds d'archives de ces institutions ne sont pas rangés dans des magasins conformes aux normes d'archivage et aux règles d'hygiène et de sécurité. Uniquement l'IVD a prévu un local doté de moyens de sécurité (chambre forte, accès avec empreinte digitale, des caméras de surveillance) pour protéger les dossiers et les enregistrements des témoignages des victimes.

10. Le traitement des fonds d'archives sur les violations des droits de l'Homme

Le traitement des documents d'archives passe par différentes opérations techniques allant de la création jusqu'à la conservation en passant par la description physique et intellectuelle en vue d'en faciliter le repérage et l'exploitation.

a. Les critères d'évaluation et de sélection des archives

La question qui se pose, après l'identification des archives sur les violations des droits de l'Homme dans les différentes structures, comment ces structures procèdent-elles pour sélectionner les documents à conserver ? Une fois le tri est fait quels moyens et outils à utiliser pour traiter ces documents ?

Plusieurs critères ont été identifiés par les instances et la société civile pour pouvoir sélectionner les documents qui vont servir plus tard comme une source de qualité offrant des usages multiples aux chercheurs, comme l'indique le tableau suivant :

Instances	Critères de sélection des dossiers
ANT	- Les archives publiques - Archives créées après la révolution
IVD	- Archives publiques - Archives privées
CSDHLF	- Archives privées - Dans la spécialité ayant une relation avec les violations des DH
CNIDV	- Archives privées - Crédibilité des personnes

	- Recevabilité des plaintes
LTDH	- Archives privées
ATFD	- Archives privées

Tableau 3 : Les critères de sélection des fonds d'archives sur les violations des droits de l'Homme

Nous avons constaté que les six responsables interviewés se servent d'un critère commun celui «d'archives privées» qui leur permettent d'évaluer la valeur d'un document qui constituera après traitement un patrimoine historique essentiel. Ce critère est important étant donné qu'il permet de distinguer la qualité du producteur de documents. La loi n°95-88 du 02 août 1988 dans l'article 23, qualifie les archives privées d'archives historiques. La même loi stipule que : « peuvent être classées par décret comme archives historiques toutes archives privées qui pour des raisons historiques présentent un intérêt public ». Ainsi, les archives sur les violations des droits de l'Homme et provenant des instances et des associations et représentant des contenus clés présentent un intérêt public puisqu'elles influent sur la vie des personnes. En plus de ce critère, nous avons remarqué que l'ANT distingue ces archives en les qualifiant comme des «*Archives créées après la révolution* ». Selon le directeur général de l'ANT, ces archives sont considérées comme des « *archives exceptionnelles* ». L'ANT n'a pas encore pris du recul pour les évaluer « tout est important pour le moment ».

b. Les moyens et procédures de traitement des fonds d'archives sur les violations des droits de l'Homme

Le traitement descriptif et intellectuel des archives sur les violations des droits de l'Homme est primordial étant donné leur caractère informatif et leur valeur de preuve pour questionner le passé, dénoncer des violations et se doter d'instruments pour demander des dommages-intérêts. En effet, ce que nous avons détecté, en faisant l'état des lieux de la gestion de ces archives par les instances et la société civile, c'est que la société civile (LTDH et l'ATFD) n'applique pas les opérations archivistiques exigées par les normes internationales à l'exception de l'ANT et

l'IVD qui utilisent les ISAD(G)¹¹⁶, ISAAR¹¹⁷ et ISO¹¹⁸ 15489. En ce qui concerne la CSDHLF et la CNIDV, elles se sont servies du système d'information et de documentation sur les droits de l'Homme, *HURIDOCS*. C'est un système qui fournit les outils et les méthodes (formats et standards sur les violations des droits de l'Homme) pour enregistrer et échanger l'information sur les droits de l'Homme. «*HURIDOCS (Systèmes d'information et de documentation sur les droits de l'Homme, international) est un réseau mondial constitué d'organisations œuvrant pour les droits de l'Homme. Il a été créé en 1982, afin de réaliser un «système universel de communication des informations sur les droits de l'Homme», [...]. En pratique, HURIDOCS s'efforce d'améliorer l'accès aux informations publiques sur les droits de l'Homme et leur diffusion, entre autres en fournissant des outils et des formations appropriés*»¹¹⁹. Le responsable de CSDHLF interviewé nous a indiqué que HURIDOCS a été traduit et adapté au contexte du monde arabe pour pouvoir élargir, ultérieurement, son usage par les organisations œuvrant pour la défense des droits de l'Homme. Toutefois, la LTDH et l'ATFD ne servent pas de méthodes et techniques de traitement appropriés et compatibles à leurs activités. Pourtant, HURIDOCS a veillé à l'adaptation de ces formats aux besoins spécifiques des organisations ayant intérêt aux droits l'Homme tels que l'intérêt porté à la surveillance des violations commises sur la femme. Champ qui peut rendre service à l'ATFD.

La description des documents se présente de différentes manières telle que définie dans le tableau suivant :

Instances	Moyens humains	Description des fonds
ANT	Personnel qualifié : des archivistes	- Traitement matériel des documents (les mettre dans des boîtes neutres)

¹¹⁶ISAD(G) (Norme générale et internationale de description en archivistique)

¹¹⁷ ISAAR (La Norme internationale sur les notices d'autorité archivistiques relatives aux collectivités, aux personnes et aux familles (CPF))

¹¹⁸ ISO 15489 (Norme internationale Records management)

¹¹⁹ Dueck, J., Guzman, M. et Verstappen, B. : *Formats standards HURIDOCS pour les évènements : un outil pour la documentation sur les violations des droits de l'Homme*, Suisse, 2002, secrétariat HURIDOC, p.1

		<ul style="list-style-type: none"> - Application des opérations de dépoussiérage, restauration des documents, enlèvement des moisissures et des petites taches - Description des documents selon les normes
IVD	Personnel qualifié : des archivistes	<ul style="list-style-type: none"> - En relation avec le contenu des plaintes : nature de l'acte, adresse de la prison, durée de la sanction, conditions de détention, nature de la torture, l'endroit, les modes de la torture. - Selon des procédures scientifiques
CSDHLF	Pas d'archiviste	<ul style="list-style-type: none"> - Nature de la violation - Nom et prénom de la victime - Date - Lieu - etc.
CNIDV	Pas d'archiviste	<ul style="list-style-type: none"> - La nature de la plainte - Par catégorie : blessés et décès - Deux catégories de dossiers (tranche d'âge et victime)
LTDH	Secrétaire qui tient les dossiers	<ul style="list-style-type: none"> - Nature de plainte : juridique, psychologique, professionnelle - Thématique (nature de la violence) - Nom, prénom, date de naissance, sujet, n° de dossier, adresse, boîte, l'année de la plainte, code pour le dossier - Les données saisies permettent par la suite l'analyse des informations et la rédaction des rapports. - Un tri se fait selon le sujet, le porteur de plainte et la date
ATFD	Archiviste contractuel pour traiter le fonds	<ul style="list-style-type: none"> - Typologie des violences (morale, physique, sexuelle, etc.) - Selon des thèmes : centre d'écoute et d'orientation face à la violence, activités ATFD, réseau, etc. - A l'intérieur de la boîte, sujet, date, n° de boîte

Tableau 4 : Le traitement des fonds d'archives sur les violations des droits de l'Homme

Par ailleurs, nous avons noté que la majorité des structures interviewées utilisent un système de classement qui leur facilite l'accès aux dossiers, à savoir le classement chronologique et

géographique. Certes, l'application des fonctions et des outils archivistiques est considérée comme une obligation légale, mais c'est aussi une nécessité vitale pour toute instance publique ou privée afin de rendre facilement les documents accessibles.

11. La préservation digitale des fonds d'archives

Le problème rencontré par les instances interviewées est celui de la recherche des moyens pour financer l'acquisition de matériel et de logiciel pour assurer la préservation digitale de leurs fonds d'archives. Il s'agit d'un usage basique des technologies (Excel, bases de données sur Access) et comme l'indique le tableau ci-dessous :

Instances	Outils
ANT	<ul style="list-style-type: none"> - Logiciel - Atelier de numérisation va servir le travail de la préservation digitale une fois les fon
IVD	<ul style="list-style-type: none"> - Application IFADA pour enregistrer les témoignages des victimes - Base de données en intranet gérée par la direction des systèmes d'information et d offert aux différentes structures de l'IVD
CSDHLF	<ul style="list-style-type: none"> - Base de données
CNIDV	<ul style="list-style-type: none"> - Bases de données sur les violations
LTDH	<ul style="list-style-type: none"> - Tableau Excel pour faire entrer les données relatives aux témoignages des victimes - Bases de données simple
ATFD	<ul style="list-style-type: none"> - Excel - Scannage des documents (organisation des documents scannés par fichier à l'intérieur

Tableau 5 : L'usage des outils technologiques pour la préservation des fonds d'archives

Le contexte post-révolutionnaire exige que les différentes instances objet de l'enquête réfléchissent à la préservation digitale pour faciliter l'usage, le partage et la diffusion des documents tout en garantissant leur pérennité.

Au cours des entretiens menés auprès des responsables des instances (ANT, IVD, CSDHLF et CNIDV) et des organisations relevant de la société civile (LTDH et ATFD), nous avons constaté que les archives sur les violations des droits de

l'Homme sont d'une grande importance dans la garantie des réparations, la reconstruction de l'histoire des pays et la constitution de la mémoire collective. C'est ce qui explique l'effort fourni par les structures concernées telles que la LTDH et l'ATFD) pour avoir des fonds leur permettant de faire évoluer leur système d'archivage vers un système numérique qui garantit la protection de leurs documents. Il est à mentionné que l'ATFD a procédé, avec les moyens de bord, à la numérisation de 80 boîtes contenant des archives relevant de ses activités ainsi que la numérisation des photos et des documents audiovisuels (migration cassette vers CD et DVD). La CNIDV étant une commission *ad hoc*, ses membres ont veillé à l'enregistrement des données relatives aux plaintes déposées par les victimes, dans un premier temps, dans une base de données. Une fois son travail accompli et après avoir présenté son rapport final, la commission a directement déposé ses documents (papier, CD, DVD, etc.) aux archives nationales (ANT) qui se chargera plus tard de leur numérisation et ce, afin d'en faciliter l'usage par les personnes concernées. La CSDHLF étant à ses débuts, a prévu la numérisation des fonds d'archives collectés et traités après la sélection des documents essentiels. Quant à l'ANT, elle dispose d'un atelier de numérisation mais elle n'a pas encore procédé à la préservation numérique de ces documents essentiels. Pour sa part l'IVD a prévu, dès le départ, les moyens nécessaires pour la numérisation des dossiers de plaintes avant leur conservation dans une chambre forte.

En somme, il est recommandé d'évoluer vers la mise en place d'une plateforme collaborative regroupant l'ensemble des institutions (instances et ONG) pour permettre un meilleur échange des ressources.

12. La communication et l'accès

L'accès à l'information et le droit de savoir sont présents dans les stratégies des instances et de la société civile. Néanmoins, la communication et l'accès aux documents d'archives sur les violations des droits de l'Homme ne sont pas encore permis étant donné la sensibilité de la nature des informations contenues dans ces documents. En interviewant les responsables des quatre instances et des deux sociétés civiles qui détiennent les archives sur les violations, ils nous ont indiqué qu'ils préservent la confidentialité et l'anonymat des personnes victimes de ces violations. Actuellement, l'accès est strictement permis aux membres de chaque instance étudiée. L'IVD nous a mentionné qu'elle va proposer des recommandations en relation

avec le droit d'accès aux victimes à l'information pour connaître la vérité. Quant à la LTDH, elle nous a précisé que le passage de l'utilisation de l'interne vers l'extérieur demande un travail archivistique selon les normes exigées (réalisation d'un calendrier de conservation par exemple) pour pouvoir ouvrir l'accès au public, «les archives ne sont pas encore traitées et organisées d'une manière scientifique pour évoluer vers l'application de la loi des archives et des droits d'accès à l'information». A cet effet, ce que nous pouvons retenir, c'est qu'il n'est pas possible, vu l'état actuel des choses, d'offrir un accès aux archives sur les violations des droits de l'Homme. Rappelons même, ce que le directeur général de l'ANT nous a avancé au sujet de ce type d'archives : «ce sont des archives exceptionnelles, on n'a pas encore pris du recul pour les évaluer, les traiter et donc les mettre à la disposition du public». Pour que ces fonds puissent être ouverts au public, ils devraient être versés à l'ANT qui est l'unique organisme garantissant leur sécurité et capable d'établir les mesures visant à faciliter l'accès à ces fonds sensibles.

13. La propriété et l'hébergement

Les archives sur les violations des droits de l'Homme forment des fonds spécifiques. Après la révolution, nous avons constaté l'intérêt porté aux archives de quelques institutions répressives ayant cessées leur activité. Etant donné que ces fonds sont la propriété de l'Etat, l'ANT a veillé à les récupérer pour les protéger contre le vol et le vandalisme (RCD, ATCE, etc.). On ne peut appliquer le principe d'intégrité sur ces fonds du moment où ils ne sont pas parvenus selon les règles de versement prescrites par la loi des archives. La révolution a mis en péril la survie de quelques documents ayant trait à la répression. La CNIDV, à la fin de sa mission, ses archives sont devenues la propriété de l'Etat puisqu'elles ont été déposées aux Archives nationales. Au sujet de leur hébergement numérique, l'ANT prévoit la signature d'une convention avec le Centre National de l'Informatique (CNI) pour conserver les archives numériques.

Selon la loi organique 2013-53 du 24 décembre 2013 relative à l'instauration de la justice transitionnelle et à son organisation et à la création de l'IVD, et notamment l'article 68, «Les travaux de l'instance sont clôturés au terme de la période prévue par la loi. L'instance confie la totalité de ses documents et dossiers aux archives nationales ou à une institution de préservation de la mémoire nationale créée à cet effet». En effet, les archives de l'IVD sont publiques inaliénable et imprescriptible selon le

responsable interviewé. De leur côté, la CSDHLF, la LTDH et l'ATFD considèrent que leurs fonds, jusqu'à nos jours, leur appartiennent et sont hébergés dans leurs locaux. Selon les personnes interviewées, après l'application des règles archivistiques, ces archives seront versées aux archives nationales pour devenir la propriété de l'Etat.

14. Vers un système de gestion des connaissances en matière de violations

Devant les enjeux liés à la collecte, le traitement et la diffusion des archives spécifiques et notamment celles qui retracent les violations commises par les instances répressives sur les citoyens, une nouvelle problématique pourrait être soulevée, en plus de celle de la conservation de la mémoire informationnelle, celle de la réutilisation des savoirs et des savoir-faire contenus dans ces fonds d'archives exceptionnels. Il s'agit de bénéficier des traces papiers ou numériques afin de faire émerger des connaissances et les rendre capitalisables et partageables. Actuellement, on assiste au passage de l'ère de la gestion des processus d'archives (identification, traitement et conservation des documents) à l'ère de la mise en valeur des connaissances stockées dans des documents d'archives. En effet, avec la gestion des connaissances, les archives prennent de la valeur par l'exploitation et l'analyse des informations contenues dans les documents d'archives et pas par leur stockage. Cette nouvelle problématique nous oriente vers la proposition avancée par M.-A. Chabin de la théorie des trois statuts opposée à la théorie classique des trois âges :

- *«Le statut de production : l'information, quel que soit son support est entièrement gérée par son ou ses rédacteurs qui l'élaborent et la mettent en forme jusqu'à sa validation et sa diffusion.*
- *Le statut de trace : L'information validée, datée et figée est préservée aussi longtemps qu'elle présente un intérêt de preuve, de gestion ou de mémoire pour celui qui l'a créée et validée. Le lien organique entre l'information archivée et son auteur est ici essentiel.*
- *Le statut de **source de connaissance**. L'information n'est plus considérée d'abord pour son lien organique avec son auteur mais en tant qu'élément de connaissance pour un public, pour une communauté afin d'étudier, de*

se cultiver, de s'informer sur un thème, un événement, une histoire»¹²⁰

C'est là que réside l'importance de la refonte des systèmes documentaires qui proposent des mécanismes de traitement de documents permettant de retrouver un document selon une description classique pour les faire évoluer vers des systèmes de gestion de connaissances qui vont offrir des mécanismes d'analyse et d'interprétation permettant d'accéder au «statut de connaissance» exploitables dans la réalisation d'action.

A cet effet, il est primordial de développer un circuit de création, de mise en œuvre, de capitalisation et d'utilisation des connaissances sur les violations commises sur les personnes physiques ou morales pour permettre de comprendre le passé, de réagir dans le présent et de se préparer pour le futur. *«L'avenir ne peut se construire sur l'oubli, sur l'occultation des événements passés. Et cela vaut pour les individus, les sociétés et l'humanité. Pour comprendre la raison d'être de la démocratie et du respect des droits de l'homme, il importe de savoir et de se rappeler comment fonctionnaient les dictatures. C'est seulement en connaissant leur histoire et en se penchant sur leur passé que les pays seront à même de se choisir un avenir affranchi de la terreur et des erreurs d'autrefois»¹²¹*. Ainsi, ce circuit devrait faire l'objet d'une attention privilégiée et d'une gestion efficace.

Conclusion

La transition démocratique que vit la Tunisie depuis sa révolution de 2011 a permis de mettre en exergue le rôle des archives dans le processus de la justice transitionnelle. Le traitement des plaintes déposées à l'IVD par les victimes des violations nécessite le recours aux archives sensibles de la police politique, la présidence de la république et autres institutions

¹²⁰ Chabin, M.A. : *Archiver, après ?* Marie-Anne Chabin, Djakarta, 2007 : Chapitre 2 - Archiver ou conserver ?

<http://www.marieannechabin.fr/archiver-et-apres/2-archiver-ou-conserver/>

¹²¹ Archives des droits de l'homme du Chili, Patrimoine documentaire soumis par le Chili et recommandé à l'inscription au Registre Mémoire du monde en 2003. <http://www.unesco.org/new/fr/communication-and-information/memory-of-the-world/register/full-list-of-registered-heritage/registered-heritage-page-4/human-rights-archive-of-chile/>

répressives. Une partie de ces fonds a été déménagée aux Archives nationales afin de la mettre à la disposition de l'instance de vérité. D'autres fonds sur les violations se trouvent chez les ONG des droits de l'Homme et des libertés de la société civile. L'étude a montré que ces archives sont en péril vu les conditions de leur conservation. De point de vue technique, elles n'ont pas fait l'objet d'un traitement archivistique adéquat permettant de les identifier de manière acceptable et de les rendre accessible à la recherche. Aucune démarche de pérennisation ou de prévention n'a été envisagée pour le moment. Ceci est dû à l'absence d'une politique nationale en matière d'archives sur les violations. Une telle politique permettrait de trancher sur des questions vitales telles que la propriété, l'hébergement, l'accessibilité, la pérennité et de reconnaître leur spécificité en tant qu'archives « sensibles » qu'il faudrait préserver des dangers afin de construire une mémoire collective sur cette question délicate. L'étude préconise un système de gestion des documents et des archives sur les violations conformes aux normes et standards de Records Management et évolutif vers un système de gestion des connaissances permettant la réutilisation des savoirs et savoir-faire contenus dans ces ressources spécifiques.

Bibliographie

1. Akindes, F. : Rapport introductif n° 3 : Les transitions démocratiques à l'épreuve des faits : réflexions à partir de expériences des pays d'Afrique noire francophone». - in : Actes su Symposium international sur le bilan des pratiques de la démocratie, des droits et des libertés dans l'espace francophone, Bamako (Mali), 1er - 3 novembre 2000
2. AGRIKOLIANSKY, E. : « Usages choisis du droit: le service juridique de la ligue des droits de l'homme (1970-1990) ». Sociétés contemporaines, 2003, no 4, p. 61-84.
3. Bazin, Paule René. -«Rôle et limites des archives dans la construction de l'histoire.». *Revue Quart Monde*, N°199 - "Forger la mémoire d'un avenir commun", Année 2006 *Revue Quart Monde*
4. BREILLAT, Dominique. Libertés publiques et Droits de la Personne Humaine. 2003.
5. Chabin, M.A. : *Archiver, après ?* Marie-Anne Chabin, Djakarta, 2007 : Chapitre 2 - Archiver ou conserver ?

6. <http://www.marieannechabin.fr/archiver-et-apres/2-archiver-ou-conserver/>
7. Chabin, M. A. « Document trace et document source. La technologie numérique change-t-elle la notion de document? », *Revue I3*. 4(1) :141-158. 05 juillet 2004, <http://archivesic.ccsd.cnrs.fr/>
8. Conseil International des Archives : *Principes relatifs à l'accès aux archives établis par le Conseil International des Archives*, adoptés par l'AGM le 24 mai 2012.
9. Couture, C. ; Lajeunesse, M. : *L'Archivistique à l'ère du numérique : les éléments fondamentaux de la discipline.*- MDullin, Sabine.- Le secret et ses usages : les archives « sensibles » en URSS. P.189. - <http://books.openedition.org/éditions-CNRS/1520?lang=fr> ontréal : PUQ, 2014
10. Dueck, J., Guzman, M. et Verstappen, B.- *Formats standards HURIDOCs pour les événements : un outil pour la documentation sur les violations des droits de l'Homme*, Suisse, 2002, secrétariat HURIDOC.
11. Georgina, Sánchez López. *Les chemins incertains de la démocratie en Amérique latine*, Ed. L'Harmattan, Paris 1993
12. Hachad, F.; Ferchichi, W., dir. : *Révolution tunisienne et défis sécuritaire. T1 : Les Archives de la dictature entre justice transitionnelle et sécurité*, Sfax, Med Ali Edition (CAEU), octobre 2014.
13. INTERNATIONAL CRISIS GROUP : *Tunisie : justice transitionnelle et lutte contre la corruption : rapport Moyen orient/ Afrique du nord*, n° 168, mai 2006
14. Linz, Juan J., « Democratic transitions and democratic consolidation », *Mimeo*, July 1991
15. UNESCO et Archives nationales de Tunisie : *Les Archives et le droit de savoir : actes de la journée de réflexion organisée à l'occasion de la journée nationale des archives par l'UNESCO et le Archives nationale de Tunisie (10 déc 2012)*, Maroc, UNESCO, 2013.
16. Quintana, A. G : *Politiques archivistiques pour la défense des droits de l'Homme*, Paris, Conseil international des archives, 2009
17. Rollinde, M. : *Le mouvement marocain des droits de l'homme: entre consensus national et engagement citoyen*. KARTHALA Editions, 2002.
18. Saleh, I. ; Bouhai, N. ; Achour, H.- *Les frontières numériques*. Paris : L'Harmattan, 2014.- 252 p.

19. Saidi, M., Communication des archives définitives,
www.archivesdefrance.culture.gouv.fr/static/2986

An Interdisciplinary Approach for Cultural Heritage Valorisation and Visualization

Basma Makhoulf Shabou^a

Maria Sokhn^b

Alex Olivieri^b

^a University of Applied Sciences of Western Switzerland,
Geneva-Switzerland

^b University of Applied Sciences of Western Switzerland,
Valais-Switzerland

Abstract

City-Zen is an interactive temporal knowledge-browsing platform that aims to valorise cultural heritages. While many organizations propose relevant data sets, they are hardly accessed, analysed and reused because of the formats inconsistency and the inappropriate information browsing and visualization. The goal of the project is to valorise the existing cultural heritage through a citizen centric design platform. The use case of this project involves a user willing to discover the history of a region and to embark in a cultural journey in the past. This project addresses a novel approach of data analysis of both assessment methods of data quality and spatio-temporal information. The proposed paper will describe the approach, its model and its implementation tests in the context of the State of Wallis.

Introduction

Nowadays multiple data and information sources are accessible and the simple fact of having to access, combine and analyse

these data requires companies to have handled several general s, models and data formats. Consequently, these last years the problem of data integration has seen a large number of scientific lean on (Sokhn, 2011). Integrating data from multiple heterogeneous sources leads to the coexistence of different data models and scheme and therefore different query languages. This situation raises the issue of heterogeneous data integration, management and retrieval. To deal with this issue, classical approaches exists such as wrapper mediator architectures where the wrappers are responsible to adapt local data such that they adhere to the common data model once exported and the mediator stores a global integrated schema and its relationships to the local scheme exported by the wrappers and recently approaches based on peer data management where no common global schema is required where the emphasis goes to the data co-existence providing base functionality over all data sources, regardless of how they are integrated (Mauroux, 2006; Aberer, 2011) have been proposed. This paper describes an original approach where our platform provide a semantic web oriented data model which acts as central repository and that deals with cases of incorrect and incomplete information by using the social web through active users engagement and data qualities metrics. In the following sections first, issues will be discussed; second, methods will be explained; third, platform will be described; fifth, data qualities metrics are detailed.

Issues

A driving force behind the development of novel services tailored to the needs of citizens is the increasing flexibility of social time. This has resulted, in the tourism area, in a stagnation of the number of long stay holidays while, at the same time, the number and relevance of short stays has grown significantly. Unlike long stays, which are more focused on the binomial sea-mountain, short stays focus on rural towns pair space (Cuvelier, 1998; Cazes, 2001). These new tourist flows are major resources for regional development, particularly in urban areas. In contrast to conventional tourist stays and interest, these short stays are no longer restricted anymore to only religious buildings, political or military points. Rather, the entire neighbourhoods of tourist destinations are now the subject of beautification strategies whose aim is to increase their attractiveness. Even though citizens (residents) may possess valuable information for other citizens visiting their home area as tourists, it is hardly usable

with today's information systems because of the inevitably scattered information resources. The quality of data is not enough studied and the final user do not have always the needed information with sufficient and adequate performance.

The goal of the project is to valorise the existing cultural heritage through a citizen centric design platform. The use case of this project involves a user willing to discover the history of a region and to embark in a cultural journey in the past. To date this goal, this project addresses (1) the data integration by making use of existing approaches to crawl and link accessible or user-generated content, (2) a novel approach of data analysis of both assessment methods of data quality and spatio-temporal information, (3) the data visualization by taking advantages of existing approaches of knowledge visualization.

Methods

In order to achieve the vision of City-Zen the components described above are essential. A knowledge sharing approach will be explored through : Data integration (Heterogeneous databases, Knowledge linking and Data analyser), Temporal multimedia browser (Data and knowledge based queries and profile-context based push up notification), and Knowledge visualization (Personalized based knowledge visualization and mobile adapted visualization and navigation).

Even if each of the described components may require a distinct research and development approach, for the overall methodology, the “Design Science Research Methodology” (DSRM) is, used The DSRM process includes the following steps: identifying issues and motivation in order to accurately plan the implementation of our proposed solution.

Defining the goals and objectives: Given the issues and motivations defined earlier, we infer the objectives of the solution in order to answer the needs and solve the problems. The objective of this step is to define the criteria against which we will be able to evaluate the solution once it is implemented.

Designing & implementing: designing the system architecture and its components concerns this step. Some elements of the architecture can be implemented just by reusing existing knowledge and technologies, but others, and given the research questions, will need to be developed.

Testing: this step is directly related to the previous and it starts sometimes with the step of design. It concerns the design and development of the proof of concept that will demonstrate the defined use cases.

Evaluating and analysing: this step is concerned with assessing whether the proposed solution meets its defined objectives. Objectives of this step are:

- Technical: to evaluate the accuracy of the results as an objective measure, the flexibility to browse the information space and the performance of the queries.
- Human: to evaluate the usability of the system and the accuracy of the results as a human evaluation given the human mental model for the decision.
- Disseminating: the objectives of this step are to disseminate knowledge through scholarly publications. It will take the form of regular reports and deliverables as well as peer reviewed articles in conferences related to diverse domains: web of data, knowledge visualization, data integration, tourism management and innovation.

Platform Architecture

The platform will start by taking advantages of existing data coming from existing resources such as vallesiana, notrehistoire, instagram, etc. For this, we will apply existing algorithms whenever possible to gather data (as data extraction is not in the scientific focus of the City-Zen project). The platform will address the processes of converting, linking, mapping and cleaning information in order to link citizens' information/knowledge (such as photos, stories, tangible objects, etc.) to the web of data (web pages, Linked open data, social sites, etc.). This aspect is combined with the dimensions of quality applied to historical archives. It focuses mainly on the definition of quality dimensions and the methods that assure their measurement based on specific indicators and variables in the context of historical archives (Makhlouf Shabou, 2011, 2014).

The platform is oriented towards citizens and will integrate their data, as they will become available. The platform will offer adequate methods and appropriate visual interfaces in order to best answer the goal of our project. For this proposal project, we assume that the data from the citizens is available. To tackle this aspect, the project integrates a gamification approach to motivate citizens to upload their data and share their knowledge. This module will be integrated to the platform in a second phase.

The platform aims to take advantages of the activities done over this platform, to that end City-Zen aims to support their users by introducing adequate dashboards to give decision makers insights allowing them to (i) monitoring critical business processes when performance does not meet goals, (ii) managing the underlying reasons of problems, and (iii) managing processes and resources to help improving decision-making and optimizing their performance.

The sought for scientific outcome is concentrated on implementing integration and linking algorithms on cultural heritage data sets and demonstrating advanced browsing through spatial, temporal and profile based querying. The different modules that will be developed will be demonstrated through a mobile application show casing the querying methods available through a friendly user interface. A city will be chosen to play the role of pilot for this project.

Data Quality Assessment Specification: Metrics & Ranking

City-Zen's project is an application determined to develop the culture in the City of Sion in the Canton of Valais. For this project to be well structured, we have to determine the specifications for the measurement of the information quality of the data that's going to be used. The information quality criteria's nomenclature will contribute to the construction of the type of usability that the user will perceived when he's will use the City-Zen's mobile application.

Methodological Consideration

Criteria Identification

Each criteria has been identified by using different documentation related to the measurement of quality dimensions applied to electronic records and archives. Two types of

information quality criteria were determined: the general information qualities applied on information sources and the specific qualities that could be applied on information content. The assessment of this information qualities criteria were performed in the thesis *Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation* (Basma Makhoulf Shabou, 2011) and in *QADEPs's Définition et mesure des qualités des archives et documents électroniques* (Makhoulf Shabou, Mellifluo, Rey, 2013). According to this thesis and in discussion with the City-Zen's team, a selection has been established amongst these quality criteria based on electronic archiving and applied to the digital information from a mobile application. Two reasons why we use those type of information quality criteria in this City-Zen's mobile application: first, for the mobile application, we are going to use electronic data and we need to applied a type of measurement of the information quality; second, this type of measurement it has been already tested in the Canton of Valais (QADEPs, 2013) because they have been several current works and reflection about archiving and what is the valuation of information.

Scoring

The measurement of each criteria is based on a ranking system that contribute specific scores reflecting the quality level of each evaluated information quality criteria.

The different Information quality levels are:

- One to three stars : If the information is relevant, trustworthy, exploitable, representative
- Four to five stars: If the information is enhanced by a relevant source (example: Archives of Valais) and a public source (contribution of a citizen who knows on the thematic), possessed more information data about the research thematic.

Nomenclature Criteria

City-Zen's application is based on several information quality criteria. Those latest represent the general s and specifics information quality criteria that determines his usability for a tourist when he visits the town of Sion in Valais. Each criteria is assessed by a scoring from one to five stars: more the information has a high score, more the information will be trustworthy, exploitable and representative for the tourist.

Tableau 1 : General and Specific Information Qualities

General information qualities

Trustworthiness

Authenticity

Identity

Integrity

Reliability

Traceability

Completeness

Legal and administrative compliance

Historical evidence

Extensive testimony

Scarcity of evidence

Exploitability

Technical accessibility

Usability

Access effectiveness

Cognitive accessibility

Logic reparability

Comprehensiveness

Juridical accessibility

Legal and regulatory authorizations

Representativeness

Institutional context

Creator relevance

Data relevance

Socio-cultural context

Contextual scarcity

Aesthetical value

Specific information qualities

Uniqueness

Information Richness

This document will describe, justify and illustrate each criteria to explore the behavior of a tourist when he's using the City-Zen application.

In the first place, the general information qualities criteria are determined by the trustworthiness, exploitability and the representativeness of the information that the tourist is seeking in the City-Zen's application. In the second place, a description of the specific information source criteria will be put in place by describing the Uniqueness and the Information Richness of the information that the tourist is seeking.

General Information Qualities

The general information quality criteria are the Trustworthiness, the Exploitability and the Representativeness that City-Zen's application have to provide for a proper operation and use by the tourist. In this part, we will describe each general information quality criteria with their sub-criteria and a specific example of a tourist practice for each criteria. A scoring is going to be set for each criteria and sub-criteria.

Trustworthiness

The tourist will trust this City-Zen's application if this one gathered all the conditions to win his confidence, this means an authentic and reliable information that will make the tourist exploit it by using it. The Trustworthiness « refers to the ability of a document to gain the trust of the user as the preferred supporting facts source. This quality depends on the authenticity and the reliability and the durability of these qualities over time » (Makhlouf Shabou, 2011, p. 115; InterPARES 2, 2013). The tourist will take a document with authentic supporting fact sources that gain his trust over a document with no identifiable source that is difficult to know who the creator is.

Authenticity : « an authentic document is a document whose can prove that is what it pretends to be, that it was produced or received by the person who pretends to have produced or received, and it was produced or received at the moment it pretend to have been » (ISO 15489-1 :2001, p.7). For the City-Zen's application, even if the user don't have physically access to information, the mobile will provide digital documents and citizen's testimony on a proper format. That's why is important to define the trustworthiness information criteria to show the integrity and identity of the document that are diffused in the application. For example, the institution

or the citizen's testimony are safe as a source? Are they identified? Also, that is why the reliability of the document is also important to trace it, to get to know if it's a complete information and if it's legal to diffuse.

Reliability: it refers to the trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about, and is established by examining the completeness of the record's form and the amount of control exercised on the process of its creation. The records forms generated using new information technologies make increasingly difficult to determine when a record is complete and whether the controls established on its creation are either sufficient or effective for anyone to be able to assume its reliability » (InterPARES, 2005)

- Traceability : it refers to the fact of documenting all movements and all uses of a document
- Completeness: The characteristic of a record that refers to the presence within it of all the elements required by the creator and the juridical system for it to be capable of generating consequences (thematical, chronological, geographical, format of the document)
- Legal and administrative compliance

Scoring & Illustrative Example

The information received by the tourist has to be trustworthy, it has to gain the trust of this person. It has to be authentic and reliable information.

Anne is a tourist that is interested by a particular building in Sion. She wants to know:

- *When the building was built*
- *The architecture's name*
- *The history of that building*
- *If they are there other buildings with the same architecture*

With the City-Zen's platform Anne will search the building that she's looking for in the City-Zen's application by writing the name of the building. She will received several type of information with a scoring for each information.

The Trustworthiness information criteria for this tourist will be evaluated by the quality of the article, who's assessed by a score of one to five stars. Anne will received several documents describing the building.

1) One star: It's a document provided by 1 non-institutional information source with a non-complete description of the building and his history. The document is provided by a non-identifiable source like a Citizen's testimony (not professional).

2) Two stars: It's a document provided by 1 institutional information source with a non-complete description of the building and his history. Provided by citizen's authentic testimony like a historian from the region for example.

3) Three stars: It's a document provided by 1 or more institutional information source with a non-complete description of the building and some of his history. This information was collected in the metadata of the Médiathèque Valais or by a citizen's testimony. It lacks of various data like the location of the building, buildings that are similar to this one and complementary information to give more interest for the tourist.

4) Four stars: It's a document provided by 2 institutional and citizen's testimony information source. Provided by multi-source information with a complete information about the building and his history. The 2 information source are combined to make a complete document that Anne can use to learn about this history of the building. It not referenced similar building or give a GPS locator.

5) Five stars: It's a document with a complete description and history of the building. It has even information about similar buildings in Valais that Anne could visit with the GPS locator. The information is completed by a citizen's anecdote and history about the building. The information received is trustworthy and it's reliable by two authentic sources: Médiathèque Valais and a Citizen's testimony and participation.

Then, Anne will choose between these articles. More the article have a high score, more they will be trustworthy information with complementary data about the building.

Exploitability

For the City-Zen's application, the information exploitability is established by the use of a mobile by a tourist to make the research. The information is ease of use because is not a physical document but a digitalized one. The information is trustworthy because it comes from a reliable source that are Médiathèques Valais, the Archives of Valais or just a history aficionado citizen that wants to show the Valais's Heritage by making a testimony.

The **Exploitability** is the dimension that « refers to the ease of use of a document, thanks to its location, retrievability, diffusion and interpretability. The exploitability depends on three types of document accessibilities : **technical accessibility**, including physical and material needed for reading ; **legal accessibility**, including regulatory and administrative environments required for the diffusion of document ; and **cognitive accessibility** that guarantees an adequate comprehension and interpretation of document contents » (ISO 15489-1 :2001 ; Makhlouf Shabou, 2011)

The information diffused is **legally accessible** because is author's wright free (only this legal open-access type of information will be put in place for City-Zen's application). The information will be **easy to comprehend** for the tourist but technical information can be also available for the history researchers that wanted to get to know more about Sion.

Scoring & Illustrative Example

The information received by the tourist has to be exploitable, located, interpreted and diffused. It has to be a technical, legal and cognitive access.

Patrick is a tourist who's a history aficionado that wants to have specific information with attached documents for a personal research about the Valère's Castle and his history. He wants to have specific information that he will use for his future publication about Switzerland's history. Patrick's want to know:

How many documents with trustful sources can Patrick find with City-Zen?

Are the documents popularized for the information to be understanding by any fan of history? (the purpose of his research is for writing a history book)

Can he use the information found on City-Zen application for reference in his book?

If he can use City-Zen application with is iPhone and keep some of the information for printing at his PC at home?

With City-Zen's application, Patrick will received several type of documents talking about Valère's Castle and his history. The exploitability information criteria for this tourist will be evaluated by the technical, the legal and the cognitive access quality of the article, who's assessed by a score of one to five stars.

One star: The document is only available in French. The document is easy to understand for the French tourist or citizen, but the Information is accessible but non-consultable physically. Patrick will only have the possibility to read the document in his mobile.

Two stars: The document is available in French, German and Italian. The document is easy to understand for the French or German tourist or citizen. Information is accessible but non-consultable physically

Three stars: It's a document by the Archives of Valais with information about the Valère's Castle and his history but with a strict abstract without specific information, details of the history and few references. The information is exploitable, located and interpreted but it lacks of details that Patrick's wants. For example, he want's also the historic details about the « Chapitre cathédral: the Chanoine's College ».

Four stars: It's a multi-language document (7 languages). All the tourist and citizen can understand easily the information provided in the document. Information is Open access and can be consulted with a request to the institution with an identifier (download if check-in as user). Patrick will do a check-in to the institution's site in order to be registered as a user so he can download the documents that he wants.

Five stars: It's a document that combines historical specific information by the Archives of Valais and a citizen's

testimony (by a history specialist in Valais for example). The history of the Chanoine's College that lived in the Valère's Castle is well described and the citizen's testimony adds a precious details to the information. The information is exploitable, located, interpreted and diffused because is an open source document, author's wright free.

Then, Patrick will choose between these articles. More the article have a high score, more they will be exploitable information with complementary data about the history of the Valère's Castle Chanoine's College.

Representativeness

This criteria « refers to the capacity of the documents to provide a significant testimony of the institutional context in which they were created. This quality depends on two essential elements: the completeness of testimony; and the representativeness of the socio-cultural context in which these documents were created » (Makhlouf Shabou, 2011).

The **Institutional Context** provides the creator or the Data relevance of the document presented in the City-Zen's application. The document can be provided by the Tourism Office for example, who's relevant because of the integrity of this institution.

For the *Representativeness*, the **Socio-cultural context** is also relevant because of the Contextual scarcity or the aesthetical value of the document presented in the application. But we can say that the aesthetical value will be assigned for the presentation of the digital document in the application.

Scoring & Illustrative Example

Sienna is a tourist interested in Sports. She wants to go to Valais to find some outdoor activities like hiking on the mountains or by a river. She wants to know all the good places like the different types of « Bisses » that are displaced in Valais (The Bisse is In the Valais, is a long canal bringing water for irrigation). City-Zen's application will put in place a choice of possible destinations to Sienna.

The information has to be representative, which means relevant and scarce, determined by the institutional and socio-cultural context.

Sienna will received information about the different types of sport courses in Valais who's updated every year by the Office of Tourism. The Socio-cultural context is Médiathèques Valais and Citizen's testimony validated by an institutional context to be representative of this institution, for example the Office of Tourism in Valais. The information can be consulted.

One star: The information received is recommended by Office of Tourism but the information displayed is not up to date (of last year for example). The information can still be useful but it lacks of precision. The Information is not scarce, it's displayed in a context summary (abstract).

Two stars: The information received is recommended by Office of Tourism but the information displayed is not up to date (of last year for example). The information can still be useful but it lacks of precision. The Information is not scarce, it's displayed in a context summary (abstract).

Three stars: The information received is recommended by the Office of Tourism but not enough relevant of all the courses displayed in Valais because the information was updated between 2014 and 2015. Information is not scarce, it's displayed in a context summary (abstract).

Four stars: The information received is recommended by the Office of Tourism and the document is relevant of all the courses displayed in Valais because the information was updated in 2016. Information is scarce by the exhaustiveness of the context

Five stars: recommended by the Office of Tourism and the document is relevant of all the courses displayed in Valais because the information was updated in 2016. Information is scarce by the exhaustiveness of the context

Then, Sienna will choose between these information and make a choice according to her interests. More the article have a high score, more they will be representative information, up to date and consultable. Now Sienna can choose between this types of representative information which one is convenient for her.

Specific Content Qualities

The specific content quality criteria are the Uniqueness and the Information Richness that City-Zen's application have to provide to give a rare, complete and precise information for the tourist. In this part, we will describe each specific content quality criteria and a specific example of a tourist practice for each criteria. A scoring is going to be set for each criteria.

Uniqueness

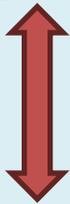
The Uniqueness data quality criteria « describes the fact that each document is related to the others within and outside the fond of which it is a part, and to the creator of the fond by a special relationship, which makes it unique. » (AAS Glossary, p.55). In short, this information quality criteria bring attention that there should be no data duplicates reported in the City-Zen's application. Each data will be unique or else the tourist will receive common and a several package of information instead of an up-to-date and exclusive information.

Asserting uniqueness of the entities within a data set implies that no entity exists more than once within the data set and that there is a key that can be used to uniquely access each entity within the data set. For example, in the City-Zen's application, each information diffused must appear once and be assigned a unique identifier that represents that information across the client applications. The dimension of Uniqueness is characterized by stating that no entity exist more than once within the data set. When there is an expectation of uniqueness, data instances should not be created if there is an existing record of that entity (Loshin, 2006).

Information Richness

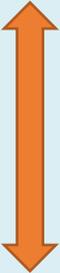
With the City-Zen's project, we discuss about the topic of information richness applied to the lowest type of rich information that is the numerical documentation (Kurstedt, 2000).

Tableau 2 : The Characteristics of media that determines richness of information (DAFT and LENGEL, 1984)

Information Richness	Medium	Feedback	Channel	Source	Language
High  Low	Face-to-face	Immediate	Developed Visual & Audio	Personal	Body, Natural
	Telephone	Fast	Audio	Personal	Natural
	Written, Personal	Slow	Limited visual	Personal	Natural
	Written, Formal	Very Slow	Limited visual	Impersonal	Natural
	Numeric, Formal	Very Slow	Limited visual	Unspecified	Numerical

The information will be formal, very slow, limited visual, impersonal and in a numeric language. However, if we use the lowest type of Information richness, for this nomenclature and the scoring we have to come out with a new table that highlights the numerical information richness. This is an approximated table explaining the highest and the lowest rich type of information applied to numerical documentation.

Tableau 3: Application of Information Richness in Numerical Documentation

Information Richness	Medium	Feedback	Channel	Source
<p>High</p>  <p>Low</p>	Numerical Index	Immediate	Developed Visual and Audio	Personal and impersonal
	Numerical Index	Fast	Visual and Audio	Personal and impersonal
	List of Key words	Slow	Visual or Audio	Impersonal
	List of key word	Slow	Limited visual	Impersonal
	Personnel word	No feedback	Limited audio	Impersonal

In this table 3, we have a new type of Information Richness applied to Numerical Documentation. The Medium defines what type of information is sent and how the information is detained in the City-Zen’s application. For example, the highest medium of Information Richness in Numerical Documentation is the Numerical Index that the tourist will receive if he’s registered in the City-Zen’s application. The lowest is for the non-registered tourist that will have only the possibility to make a research with his personal keywords. The Feedback is how the information is checked understandable and how is interpreted by the tourist. In this City-Zen’s application, we have two data sets from the Valais Wallis Digital and the Médiathèques of Valais, and for enhance this data sets, the citizens of Sion can participate making a testimony about a City-Zen’s thematic. For example, the highest feedback of Information Richness is a registered tourist who will receive an immediate feedback of

his researches and get to know immediately what the information source is. The lowest feedback for a tourist is not to have one without knowing who the information source is. The Channel is more the form of distribution of the application, if the information will be diffused on Visual or in Audio or to have the possibility to have both. For example, the registered tourist will receive a developed visual and audio of the information, making it the richest channel of numerical documentation. The lowest will be the non-registered tourist that will only receive a limited visual of the information. The Source is the type of information that the City-Zen's application have use for guiding the tourist in his researches. The highest source in Information Richness is when there is a combination of an official source (impersonal) like an information of the Tourism Office in Sion and a citizen's testimony (personal) that's not qualified as an official source. For this table, we have eliminated the Language part because we have only digital language.

This work will inspire the part of Uniqueness and Information Richness in the Specific Content Qualities shown on table 4 (Information Quality Metrics and Ranking)

Regarding the previously information quality criteria and metrics, we established a table representing the various specificities allowing to detail these criteria, who are displayed in two types of information quality criteria: the general information qualities and the specific qualities.

Tableau 4: Information Quality Metrics and Ranking

<i>Information Quality Levels</i>					
<i>Score</i>	*	**	***	****	*****
<i>Criteria</i>					
<i>General Information Qualities applied on sources</i>					
<i>Trustworthiness</i>					

<i>Reliability</i>	Provide d by 1 non-institutional information source Non-complete description	Provide d by 1 institutional information source Non-complete description	Provide d by 1 or more institutional information source Non-complete description	Provided by 2 institutional and citizen's testimony information source	Provided by institution and professional testimony sources with exhaustive description
<i>Authenticity</i>	Provide d by source non-identified	Provide d by Citizen's authentic testimony	Provide d by Authentic professional testimony	Provided by multi-source information	Provided by authentic multi-source information
<i>Exploitability</i>					
<i>Cognitive Accessibility</i>	French	French/German/Italian	French/German/Italian/English/Spanish	Multi-language (7 languages)	Multi-language (more than 10 languages)
<i>Technical Accessibility</i>	Easy to understand for the French tourist or citizen	Easy to understand for the French/German/Italian tourist or citizen	Easy to understand for the tourist or citizen	All tourists and citizens can understand easily	All tourist and citizen can understand easily
<i>Juridical Accessibility</i>	Information is accessible	Information is accessible	Information is accessible	Information is Open	Information is Open access and

	le but non-consultable physically	le but non-consultable physically	le but non-consultable physically	access and can be consulted with a request to the institution with an identifier (download if check-in as user)	can be consulted with a request to the institution with an identifier (download if check-in as user)
--	-----------------------------------	-----------------------------------	-----------------------------------	---	--

Representativeness

<i>Institutional Context</i>	Recommended by Office of Tourism but information not up to date (of last year)	Recommended by Office of Tourism but information not up to date (of last year)	Recommended by Office of Tourism but information not up to date (of last year)	Recommended by Office of Tourism and information is up to date	Recommended by Office of Tourism and information is up to date
------------------------------	--	--	--	--	--

<i>Socio-cultural Context</i>	Information is not scarce, in a context summary (abstract)	Information is not scarce, in a context summary (abstract)	Information is not scarce, in a context summary (abstract)	Information scarce by the exhaustiveness of the context	Information scarce by the exhaustiveness of the context
-------------------------------	--	--	--	---	---

Specific information Qualities applied on contents

<i>Score</i>	*	**	***	****	*****
--------------	---	----	-----	------	-------

<i>Criteria</i>					
<i>Uniqueness</i>	-Density of information, lacks of precision, - Abundance of format, media, information typology, subject, period / context	-Density of information, lacks of precision, - Abundance of format, media, information typology, subject, period / context	- Selection of information, lacks of precision, -A selection of format, media, information typology, subject, period / context	- Exclusive Information, precise and indexed - Scarcity of format, media information typology subject, period/context.	-Exclusive Information , precise and indexed - Scarcity of format, media information typology subject, period/context.

<i>Information Richness</i>	The Information will be searched by personal words, there are no feedback available, the information will be limited audio, with an Impersonal source (ex. Tourism Office)	The information will be searched by a list of keywords, the feedback will be slow but available, the information will be limited in Visual, with an Impersonal source (ex. Tourism Office)	The information will be searched by a list of Keywords, the Feedback will be slow, the information will be diffused in Visual or in Audio, with an Impersonal source	The information will be searched by a Numerical Index, the Feedback will be Fast, the information will be diffused in Visual or in Audio, with a Personal and an Impersonal source (ex. Citizen's testimony and Tourism Office)	The information will be searched by a Numerical Index, the Feedback will be Immediate, the information will be diffused with a developed Visual and Audio, with a Personal and an Impersonal source (ex. Citizen's testimony and Tourism Office)
-----------------------------	--	--	--	---	--

The platform proposed in this project will take advantage from this work and will propose the partial automation of detailed metrics.

Conclusions

The main expected outputs of our research are:

- The centralization and the classification of heterogeneous data derived from public and private records,

- The reduction of the access time to spot information in the city,
- The valorization of the private records able to promote the recognition of the contributors as participating to a community of interest and to the creation of a story, fostering the sharing of information,
- The analysis of the profile and the behavior of the users (space, time, content) in order to support public decision-makers (e.g. local heritage enhancement, town planning, etc.).

A fundamental limit to the private records addition system is based on the control of the classification and of the pertinence of the added data. Gamification, able to encourage and to reward contributors, should also allow penalizing excessive uses.

Beyond the technical developments, further researches will focus, in the one hand, on the discovery and the addition of new reliable information sources, and in the other hand, on the use cases of private users and decision-makers.

References

Aberer, K (2011). Peer-to-Peer Data Management. . Synthesis Lectures on Data Management, Morgan & Claypool Publishers 2011.

Barrow, T. (2010). Information Richness Theory. *Time Barrow Complaining Digital Orality* [on ligne] : <http://blog.timebarrow.com/2010/11/information-richness/>

Cazes, G (2001). « L'émergence d'un nouveau système vacancier : temporalités et territorialités en mutation », *Hommes et Terres du Nord*, n°2, pp. 63-70.

Cuvelier, 1998 CUEVELIER Pascal, 1998, Anciennes et nouvelles formes de tourisme. Une approche socio-économique, Paris : L'Harmattan.

Daft, R. L., & Lengel, R. H. (1984). Information Richness – a New Approach to Managerial Behavior and Organizational Design. *Research in Organizational Behavior*. Avril 1983. Vol. 6, pp. 191-233.

Instagram, capture and share the world's moments,
<https://instagram.com/>

INTERPARES TRUST (2001). *The InterPARES Glossary* [en ligne]. Vancouver: International Research on Permanent Authentic Record in Electronic System (InterPARES). [on ligne] :
http://www.interpares.org/book/interpares_book_q_gloss.pdf

Kurstedt, H (2000). Define Information Richness. In: *Management System Theory Applications and Design* [on ligne] : <http://www.wmich.edu/emrl/vt/pdf/Book/4015/1.4.3-1.4.5.8.pdf>

Lee, B (2005). Authenticity, Accuracy and Reliability: Reconciling Arts-related and Archival Littérature. *InterPARES 2 Project*. [on ligne] :http://www.interpares.org/display_file.cfm?doc=ip2_aar_arts_lee.pdf

Loshin, D (2006). *Monitoring Data Quality Performance Using Data Quality Metrics*. [on ligne] :
https://it.ojp.gov/documents/informatica_whitepaper_monitoring_dq_using_metrics.pdf

Makhlouf Shabou, (2011a), Measuring the Quality of Records to Improve Institutional Memory. IEEE International Professional Communication Conference - Cincinnati, 2011.

Makhlouf Shabou, (2014), Le projet QADEPs : un outil au service de la pérennisation des archives publiques. Dans Hiraux, Françoise & Mirguet, Françoise, *De la préservation à la conservation : stratégies pratiques d'archivage*. Louvain-la-Neuve : Academia l'Harmattan, 2014, pp. 87-98.

Makhlouf Shabou, (2011b). Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation [en ligne]. Montréal : École de bibliothéconomie et des sciences de l'information. Thèse de doctorat. [on ligne] : <http://hdl.handle.net/1866/4955>

Makhlouf Shabou, (2015). Digital diplomatics and measurement of electronic public data qualities: What lessons should be learned? *Records Management Journal* [en ligne]. Vol. 25 Issue: 1, pp. 56 - 77. [on ligne] :<http://www.emeraldinsight.com/doi/full/10.1108/RMJ-01-2015-0006>

Makhlouf Shabou, B; Mellifluo, L & Rey, R (2013). *QADEPs: définition et mesure des qualités des archives et documents électroniques*. Genève : Haute école de gestion [Consulté le 10 mai 2016]. Rapport de recherche

Mauroux, P (2006). Emergent semantics: rethinking interoperability for large scale decentralized information systems. 2006

Notrehistoire, Partagez vos archives et contribuez à illustrer l'histoire de la Suisse, <http://www.notrehistoire.ch/>

Organisation mondiale de normalisation (2001). *Information et documentation : Records Management*. Genève: ISO. Norme internationale ISO 15489-1: 2001

Society of American Archivists (2005). *A Glossary of Archival and Records Terminology* [en ligne]. Chicago: Society of American Archivists. [on ligne] : <http://files.archivists.org/pubs/free/SAA-Glossary-2005.pdf>

Sokhn, Maria. (2011). Conference advanced multimedia information management and retrieval (PhD thesis, telecom Paristech)

Tarata, C (2015). Data quality dimensions – from Accuracy to Uniqueness. *Performance Magazine* [on

ligne] <http://www.performancemagazine.org/data-quality-dimensions-from-accuracy-to-uniqueness/>

Tieche, J (2015). *La mesure des dimensions de la qualité des archives électroniques : apport des textes normatifs en matière d'archivage électronique à long terme* [en ligne]. Genève : Haute Ecole de Gestion. Travail de Bachelor. [Consulté le 5.04.2016]. Disponible à l'adresse : http://doc.rero.ch/record/258018/files/TDB_Tieche_Julien.pdf

Vallesiana, le patrimoine numérique du valais, <http://www.vallesiana.ch/>

Expériences de numérisation à la Télévision Tunisienne : processus et défis Etude de cas : les supports type 02 pouces

Kaouthar HAMED

Unité Archives Télévisées-Télévision Tunisienne

Introduction :

Instaurer le Droit du libre accès à l'information des organismes publics en Tunisie est un processus en voie de développement, mis à l'examen une série de textes réglementaires référentiels le prouve : un décret -loi en 2011, un premier texte d'application de la constitution du Janvier 2014, une loi organique en 2016.

Le corpus législatif en la matière introduit les composantes du processus sus indiqué, nécessaire et insuffisant, d'autres composantes tel que les ressources humaines qualifiées, les ressources matérielles adéquates, réviser le statut des documents administratifs, définir les instances publiques concernées, qualifier les chargés d'accès à l'information et les tiers, afin de paramétrer l'accès et la publication de l'information suivant des dispositifs prédéfinis.

Unir les ressources humaines et matérielles sous l'emblème du corpus législatif et suivre leurs intégrations ; mesurer l'efficacité voire l'efficience du système d'information sont des préoccupations majeures pour les professionnels de l'information.

Le secteur public est focalisé sur les missions essentielles d'information, d'éducation et de divertissement, réunir autant de vocations garantie la survie des organismes et institutions publiques.

En ce qui concerne le secteur public de l'audiovisuel, représenté par les canaux de la Radio Tunisienne et de la Télévision

Tunisienne aucune spécification légale mentionnée, c'est l'opacité qualifiée.

Partant de cet examen et de ce dernier constat, l'information née sous la forme audiovisuelle est exclue du dispositif législatif ; elle est destinée à la diffusion à travers les canaux sus indiqués, en prenant l'exemple de la Télévision Tunisienne qui détient un gisement d'images animées estimé à plus de 280.000 heures de diffusion ,un fonds original unique en genre et en nombre depuis un demi-siècle avec une première chaîne généraliste et une autre destinée à la jeunesse ,une interrogation resurgit : quelles sont les mesures adéquates pour pérenniser l'accès et la diffusion des archives télévisées qui ont marqué l'histoire de l'audiovisuel tunisien ?

Cet article vise à répondre à cette question centrale en exposant les différents projets qui ont été mis en œuvre sur la voie de la pérennisation de l'accès au fonds télévisuel, les moyens alloués et les entraves rencontrés sont à considérer, l'expérience Tunisienne est semblable à celle d'autres institutions publiques du secteur de l'audiovisuel dans le Monde arabe et des pays du pourtour méditerranéen.

1- La Télévision Tunisienne : la mémoire vive de l'image animée

Fondée le 31 Mai 1966, la Télévision Tunisienne est l'institution publique qui détient le fonds télévisuel le plus riche et le plus diversifié depuis cinq décennies.

C'est grâce à l'ingéniosité des hommes de Lettres, des cinéastes, techniciens et hommes de théâtre et artistes de tout genres que la Télévision Tunisienne a produit des émissions phares des années 60 et 70 à l'instar des séries : *oumi traki,haj klouf,mhal chahed,les contes d'El Aroui, janet el atfel ,mathaf el aghani*¹²²et bien d'autres délices télévisuels (visites des ténors arabes en Tunisie, soirée festives ,documentaires, discours politiques ,galas et festivals.

Associée à la Radio Tunisienne depuis sa création, la Télévision Tunisienne fut séparée administrativement en 2007, elle devient une institution autonome, et reste sur le plan artistique la source

¹²² Les références concernent des anciennes productions télévisées :
أمي تراكي،الحاج كلوف،محل شاهد،حكايات العروي،جنة الاطفال،متحف الأغاني

des créations dramatiques, des émissions informatives, des programmes de divertissement, de jeux et de compétitions retransmis en direct.

Les défis étaient d'assurer un service public d'information, de sensibilisation et d'apprentissage, les principales missions de la télévision publique (2) ¹²³sont :

- la contribution à la promotion de la communication, de la culture et de la politique générale de l'Etat tout en se référant aux attributs de l'identité nationale,
- l'enrichissement et le développement du paysage télévisuel au niveau de la production et de la diffusion,
- l'enrichissement du contenu de la communication télévisuelle à travers le suivi et la fourniture de l'information au niveau national et régional, la facilitation de l'accès à l'information et la couverture des événements à l'intérieur et à l'extérieur du pays,
- la production et la commercialisation des travaux audiovisuels et les spots télévisuels, la vente des espaces publicitaire et la réalisation des opérations de sponsoring et de parrainage.
- Vendre tout produit et toute prestation de services en rapport avec les activités audiovisuelles.
- l'achat des droits des produits en vue de leur diffusion ou de leur vente,
- l'exploitation des espaces et des équipements de la production télévisuelle,
- la promotion de la coproduction internationale et les échanges dans le secteur de l'audiovisuel,
- la production des œuvres de fiction, les courts et les longs métrages télévisés, les séries et feuilletons et films documentaires
- l'utilisation des nouvelles technologies de production,
- la conservation et la numérisation des archives audiovisuelles,

¹²³ Décret 2007-1868 du 27 juillet 2007 portant sur création, organisation administrative et financière et modalités de fonctionnement de la Télévision Tunisienne

2- Œuvre ou production ou archive conjugués au mode audiovisuel

Uniques, irremplaçables et intemporelles, telles sont les qualificatifs intrinsèques des œuvres télévisées, de 1966 à 2005 date de l'avènement de la première chaîne privée en Tunisie "Hannibal tv", la télévision tunisienne fut la référence des œuvres télévisées produites et archivées par les soins des professionnels de l'époque.

Se référer à la législation tunisienne pour définir les archives audiovisuelle n'est pas une option concluante, dans son article premier la loi n° 88-95 du 02 Août 1988 relative aux archives donne une définition exhaustive et non spécifique aux archives audiovisuelles.¹²⁴

Les textes de lois sur les documents électroniques renvoient au commerce électronique et à des domaines d'activités loin de la gestion des documents quelque soient leur supports et forme soulignées dans la loi relative aux archives.

Les arrêtés et circulaires prônent la gestion du format texte, les archives audiovisuelles n'y figurent pas, ce sont des archives indéfinies au sens juridique, c'est l'usage, la production et la diffusion qui les révèlent.

Helen P.Harrison définit le document audiovisuel comme étant pièce formant le patrimoine audiovisuel et applique l'appellation document audiovisuel à «tous les enregistrements visuels (avec ou sans la bande son) et à tous les enregistrements sonores, quels que soient le support matériel ou le procédé d'enregistrement ». (**Harrison, 1992**)¹²⁵

Marie-Anne Chabin pense que les archives audiovisuelles ont « d'abord l'avantage de l'attractivité, au travers du son et de l'image qui captent l'attention sans que l'on ait à produire l'effort de lire et de comprendre un texte....Dans une société où prime l'audiovisuel au quotidien ,la forme audiovisuel d'une

¹²⁴ « les archives sont l'ensemble des documents quelque soient leur date, leur forme et leur support matériel produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé, dans l'exercice de leur activité ».article 01 loi n° 88-95 du 02 Août 1988 relative aux archives, JORT 52.

¹²⁵ (**Harrison, 1992**) « les archives audiovisuelles à travers le monde chapitre 14,chapter 14 audiovisual worldwide » / Helen P.Harrison, UNESCO,1992,p194.

archive est un atout pour l'utilisateur ,qu'il s'agisse du consommateur final qui a le choix entre l'image sonorisée et le texte, ou qu'il s'agisse de l'utilisateur-producteur qui exploite les archives pour une nouvelle production ».(Chabin, 2014)¹²⁶

Il n'est pas aisé de définir les archives audiovisuelles, qui intègrent les notions de la production, de l'audiovisuel, de l'œuvre, du support et du contenu ; toutefois parler de l'œuvre audiovisuelle rime avec ses qualificatifs inhérents.

Wahid Gdoura pense que les archives audiovisuelles constituent des témoignages multiples : historiques, culturels et sociaux pour les événements passés et avenir. (Gdoura, 2010)¹²⁷

Mohamed Almanasir déclare que les archives audiovisuelles ont une valeur permanente à l'heure où la transmission des informations est sans frontières. (Al Manassir, 2010)¹²⁸

3-Les événements créent les œuvres et le numérique recrée le patrimoine

La valeur probatoire des œuvres audiovisuelles provient des fonctions qu'elles représentent : la compréhension, la mémorisation, le témoignage et de valorisation.

Longtemps contestée à cause des effets additifs de la technologie et des possibilités techniques d'ajout et de suppression, les œuvres audiovisuelles sont des preuves tangibles qui s'associent au registre des preuves, d'attestation de droit au format papier.

En Tunisie, les œuvres audiovisuelles, précisément celles télévisées marquent toutes les phases du développement du pays, les reportages télévisés sont des narrations des premières campagnes de vaccination, les campagnes pionnières du planning familial, de l'école publique pour tous, de l'ascension

¹²⁶ (Chabin, 2014) qu'est ce qu'une archive audiovisuelle ?/Marie-Anne Chabin, E-dossier de l'audiovisuel : L'Extension des usages de l'archive audiovisuelle,France, INA, 2014, p3.

¹²⁷ (Gdoura, 2010) "القوى العاملة في مجال الأرشيف السمعي البصري: الكفاءات و المؤهلات" /Wahid Gdoura, revue ASBU,n°4,2010,p58.

¹²⁸ (Al Manassir, 2010) الحفظ و الصيانة لأرشيفات الصوت و الصورة /Mohamed AL Manassir, revue ASBU,n°4,2010,p38.

des femmes à l'éducation, de la communication politique et publique par les discours des leaders politiques en direct et en différé.

Le processus de la numérisation requière l'existence d'un capital humain suffisant en nombre et en qualification, d'un budget alloué et grandissant considérant l'obsolescence des supports, en fait le progrès technologique liés aux médias est évolutif, des supports apparaissent et d'autres s'anéantissent, l'accès aux fonds télévisuels dépend de la maîtrise des technologies nouvelles. Ainsi **Dominique Saintville** affirme que « chaque institution apporte des réponses souvent similaires sur le plan purement technique, mais très diversifiées dans leur mise en œuvre ». (**Saintville, 2009**)¹²⁹

Les sociétés qui opèrent dans le secteur de la numérisation des œuvres télévisées (restauration et transfert) ne sont pas nombreuses, elles gardent un savoir faire et des lecteurs révolus et les relie à des mécanismes à la pointe de la technologie.

Outre la nécessité d'un Budget consacré à la numérisation, un capital matériel est requis pour achever la numérisation : des équipements tel que (lecteurs, matériels de maintenance, machines enregistreuses, dépoussiéreuses....) en plus d'un capital humain suffisant en qualification et en effectif.

En Tunisie, deux ans après l'instauration de la Télévision Tunisienne, l'Unité des archives télévisées, appelée à l'époque (Mai 1968) *filmothèque*, par rapport aux supports de conservation de l'époque, les télécinémas (supports films), a été le premier noyau de la documentation audiovisuelle.

Cette cellule de stockage des journaux télévisées initialement, s'est agrandi ,des services de documentation, d'archivage ,de conservation et de numérisation des archives télévisées sont nés afin d'ordonnancer les joyaux télévisuels ,qui furent restaurés et numérisés depuis peu.

Les expériences de numérisation des œuvres télévisées s'insèrent dans ce cadre de coopération professionnelle entre les organismes publics de l'audiovisuel.

¹²⁹ (Saintville, 2009) »la stratégie de sauvegarde et de numérisation des archives de l'Institut National de l'Audiovisuel, 1999-2015 »/Dominique Saintville, the international preservation news, IFLA, n°47, 2009,p18.

La Télévision Tunisienne ratifie les conventions de partenariat avec l'institut national de l'audiovisuel (**INA**), elle est membre de la Fédération Internationale des Archives Télévisées (**FIAT**), de la conférence permanente de l'audiovisuel méditerranéen (**COPEAM**), en tant que diffuseur elle est également membre de l'union des diffuseurs des pays arabes (**ASBU**) que la Tunisie abrite son siège. Ces alliances favorisent les opportunités d'échange d'expertise, des équipements et des expériences avec les différents diffuseurs du bassin de la méditerranée et des pays arabes.

4-L'expérience Cap Med (2000-2005)

La Télévision Tunisienne a toujours prôné l'accès aux archives tout en respectant les règlements en vigueur, le corpus légal concernant les archives audiovisuelles est mince et ferait l'objet d'un article spécifique.

Le partenariat avec les organismes spécialisés tel que l'Institut national de l'audiovisuel (INA-France) fut une voie privilégiée pour déclencher les phases de restauration curative et de numérisation pérenne.

Dotée d'un capital humain composé par des techniciens, des ingénieurs spécialisés en la matière, la Télévision Tunisienne s'est insérée dans ce processus de numérisation depuis l'année 2000, en ratifiant des accords de partenariat avec l'institut national de l'audiovisuel, le projet **Cap Med** qui s'est étalé de 2000 à 2005, est un fruit de ce partenariat.¹³⁰

Dans une des phases préalable à la numérisation, **Cap Med** a eu le mérite de tirer la sonnette d'alarme sur la disparition progressive des matériels de maintenance, les pièces de rechange, en fait la technologie vidéo analogique est révolue, l'industrie favorise depuis la fin des années 90, les opportunités d'accès aux contenus de la technologie numérique.

Initié par l'Institut national de l'audiovisuel et financé par le Programme **Euromed Audiovisuel** à hauteur de 04 millions d'euros, le projet a regroupé les chaînes de télévision publiques de plusieurs pays de la Méditerranée : l'Algérie (ENTV), Chypre

¹³⁰ Le projet CapMed est issue d'une convention signée entre l'Etablissement de la radiodiffusion télévision Tunisienne (ERTT) et l'institut national de l'audiovisuel (INA) France en 2002.

(CyBC), l'Egypte (ERTU), la Grèce (ERT), l'Italie (RAI), la Jordanie (JRTV), Malte (PBS), le Maroc (2M et TVM), l'Autorité Palestinienne (PBC), la Tunisie (ERTT) et la Turquie (TRT), ainsi que deux partenaires institutionnels, la MMSH (Maison Méditerranéenne des sciences de l'Homme) et l'IOM (Institut Hellénique de l'Audiovisuel).

Ainsi, **CapMed** est le premier réseau qui a rassemblé les archives télévisuelles des institutions publiques du bassin méditerranéen, il exprime la volonté d'améliorer la préservation du patrimoine télévisuel et d'en faciliter la circulation et l'exploitation.

Considérant la fragilité des supports et l'acuité des menaces de disparition qui pèsent sur une portion considérable des œuvres les plus anciennes.

La première étape a consisté à réparer ou remplacer les équipements de lecture nécessaires au transfert des anciennes productions télévisées sur des supports contemporains et exploitables.

Les télévisions partenaires ont eu le privilège d'acquérir des équipements pour le transfert des anciens formats vidéo (02pouces, 01Pouce, 3 /4 de Pouce), des machines de lecture des films et de machines à nettoyer.

Dans une deuxième étape la formation des techniciens d'exploitation, a permis le transfert de plusieurs milliers de documents sur des supports actuels.

Un cycle de formation a été accordé à des documentalistes afin de parfaire le traitement des données transférées et favoriser l'accès à un public plus large de professionnels et de chercheurs via l'usage des outils informatiques.

Les 13 partenaires du réseau, ont alimenté une base de données thématique, multimédia et multilingue (français, arabe, anglais), près de 4000 documents d'archives d'un grand intérêt culturel, équivalant à plus de 3000 heures de programmes sont présentés sur le site.

Les thèmes de cette base de données CapMed sont les suivants : Paysages, Histoire, Economie, Sciences, Culture, Civilisation, Modes de Vie, Politique, Sports.

5-L'expérience Med Mem (2009-2014)

Med Mem vient s'associer à l'essence du projet Cap Med, après avoir aidé les télévisions méditerranéennes à préserver et à conserver les équipements nécessaires au transfert et à former les techniciens et les documentalistes à manier les machines et les nouvelles données, vient alors la phase de la sélection des contenus représentatifs du patrimoine culturel commun des pays du bassin de la Méditerranée.¹³¹

La Télévision Tunisienne a ratifié en 2012, un accord de partenariat avec l'Ina instigatrice du projet avec le concours de la *copeam* (Conférence Permanente de l'Audiovisuel Méditerranéen). *Med-Mem* est cofinancé par l'Union européenne dans le cadre du programme *Euromed Héritage IV*.

Véritable mosaïque culturelle et historique, *Med Mem* est le plus grand portail renfermant plus que 4000 documents audiovisuels accessibles gratuitement au public en trois langues (Français, Arabe et Anglais).

Plus de 40 spécialistes ont enrichi les fiches documentaires par des éclairages et de regards croisés sur les contenus offerts dans le portail.

Vingt partenaires dont 14 télévisions du pourtour de la Méditerranée, 03 organismes professionnels et des partenaires culturels et scientifiques de haut niveau.

Dans la dernière phase du projet, un réseau de relais de diffusion s'est organisé afin de véhiculer les thématiques *Med Mem*, des Musées, des institutions de recherche, des facultés, des Bibliothèques, la pédagogie des cycles d'enseignement passent aussi par les contenus *Med Mem*.

La contribution de la Télévision Tunisienne fut l'ajout de **174** programmes en langue arabe sous titrés en français avec des notices documentaires descriptives représentant les thématiques de *Med Mem* : coutumes, traditions et paysages de la Tunisie. Pour ce qui est des relais de diffusion Tunisiens, 03 relais ont répondu à l'initiative à savoir : *L'institut supérieur de Musique de Tunis, la Bibliothèque Nationale et l'institut des Beaux Arts de Tunis*.

Les réseaux de sociabilité, **Facebook** et **twitter** se sont ralliés à la tendance *Med Mem*, en vue de rassembler un public jeune

¹³¹ La convention entre la Télévision Tunisienne (TT) et l'Institut national de l'audiovisuel (INA) fut signée en Mars 2012.

avide des connaissances et désireux de localiser historiquement et culturellement les pays du pourtour de la Méditerranée.

6-Les capitaux propres : la numérisation des supports 01 pouce C (VPR)

Outre les partenariats fructueux, la Télévision Tunisienne a misé sur ses propres ressources humaines et matérielles, c'est la volonté de préserver les fondements de la mémoire audiovisuelle collective, ainsi la totalité des supports 01 pouce, utilisé à la Télévision tunisienne du 1984 à 1996 a été transféré par les soins des techniciens de la Télévision Tunisienne soit **2671** supports équivalent **2611** heures et sont en cours d'analyse documentaire qui dévoilera les émissions et autres contenus qui n'ont pas pris une ride pour le plaisir de la programmation télévisée des deux chaînes nationales.

Une salle polyvalente pour les transferts a été installé depuis 2015 au niveau de la Télévision Tunisienne, afin de transférer l'ensemble des supports 01 pouce B (BCN) et 01 pouce C (VPR), la collecte, la réparation, mise en marche et la maintenance des lecteurs est faite par les ressources propres de la Télévision Tunisienne.

Ces missions qui sont payantes en dehors de la Télévision Tunisienne et inexistantes également quelques institutions internationales privées sont sur le créneau de la maintenance et le transfert des archives audiovisuelles, l'expertise se trouve chez les techniciens de l'image.



Figure 1- lecteur 01 pouce

BCN : le 1 pouce Type B est un format vidéo développé en 1976 par la Bosch Fernseh, une filiale de Bosch en Allemagne. Sa bande d'une largeur de 1 pouce se trouvait sur une bobine. Il ne connut pas le succès de son concurrent direct le 1 pouce Type C. www.scart.be/

VPR : Le format 1 pouce de Type C est un format vidéo professionnel sur bobine qui fut co-développé et introduit par Ampex et Sony en 1976. Il remplaça le format qui dominait à l'époque à savoir le Quadruplex 2 pouces, du fait de sa plus petite taille et de la qualité vidéo légèrement supérieur des enregistreurs. www.scart.be/



Figure 2- bande type 01 pouce

Le nouveau fonds numérisé a permis aux producteurs des chaînes nationales d'imaginer des émissions tirées du patrimoine tel que « telvezt zman »(télévision d'antan), « **saharat men archive** »(soirées d'archives) « filmothèque » et autres émissions qui utilisent partiellement ou intégralement des émissions numérisées.

Ce qui marque l'usage des supports à la télévision tunisienne est leur variété, 08 standards professionnels ont été utilisés depuis l'instauration de la diffusion télévisée (supports films, les supports 02 pouces, les supports 01 pouce (BCN/VPR), les supports ¾ de pouce, les supports Betacam analogique, les supports Betacam numérique, les supports DVC pro).

Ces différents standards impliquent une multitude de projets de numérisation adéquats aux spécificités techniques de chaque support, d'où l'allocation des ressources humaines principalement qui manient un savoir faire qui tend à disparaître.

Une nouvelle expérience de numérisation des supports $\frac{3}{4}$ de pouce Umatic va être lancée prochainement afin de sauvegarder une nouvelle parcelle de la mémoire audiovisuelle.



Figure 3- bande $\frac{3}{4}$ de pouce /U-matic

Un travail laborieux caractérise les processus de numérisation des fonds télévisés, la jonction simultanée des capitaux financiers, matériels et humains forme le défi majeur à remporter.

Tout au long de son cursus la Télévision Tunisienne a donné vie à une multitude de chaînes à savoir, la chaîne 1, **Canal 21**, **Chaîne 2 francophone**, **chaîne maghrébine** et à chaque chaîne son fonds d'archives, véritable sentier de recherches pour les producteurs, les chercheurs et les historiens de l'image.

6- Expérience de numérisation des supports 02 pouces :

Les supports 02 pouces ne sont pas les supports les plus anciens, mais ce sont les supports les plus menacés, leur numérisation concerne quelque **12000** supports que la technologie les a rendu obsolètes.

Suivant le lexique édité en octobre 2008, ces supports appartiennent à « un système d'enregistrement magnétoscopique quadruplex. Contrairement au système à balayage hélicoïdal, qui ne comprend qu'une seule tête de lecture, ce système magnétoscopique comprend quatre têtes d'enregistrement vidéo. Souvent appelé système tétraphonique, il a été conçu par **Ampex** dans le milieu des années 50. Il permettait d'obtenir une

meilleure qualité de résolution et de couleur que le système à balayage hélicoïdal. Jusqu'à tout récemment, le système quadruplex constituait la norme de radiodiffusion, car des systèmes à balayage hélicoïdal moins dispendieux ont commencé à le supplanter. »¹³²



Figure 4- bande type 02 pouce



Figure 5- lecteur 02 pouces

¹³² <http://www.marc21.ca/MaJ/BIB/B007-EnregistrementVideo.pdf>.

Utilisés à la Télévision Tunisienne depuis la fin des années 70 et jusqu'à la moitié des années 80, ils résument deux décennies riches en événements, en histoire audiovisuelle.

Outre les capitaux propres de la Télévision Tunisienne, les expériences de numérisation furent des demandes ponctuelles de transfert que la production nécessite pour la réalisation des films documentaires particulièrement et autres manifestations nationales.

La numérisation massive était une suggestion présente mais pas pressante, c'est l'urgence, la volonté et l'état des supports qui ont mené au lancement du premier appel d'offres international en 2011. Ayant pour objet la numérisation des archives audiovisuelles type 02 pouces.

Des listings regroupant les métadonnées préliminaires des contenus ont été dressés par les soins des documentalistes spécialistes en audiovisuel, répertoriées par genre pour faciliter le choix des contenus à numériser pour les besoins de la programmation, la production et les actualités.

Ce projet pionnier a connu quelques difficultés structurelles qui ont été dépassés, il a été repris en 2013, l'exécution accordée à la Société candidate VECTRACOM qui a remporté les conditions annoncés dans le cahier des charges du marché en question.

Le projet a été entamé en décembre 2014 pour aboutir en juillet 2015 à la numérisation de **1500** supports qui ont été convertis sur **3018** bandes numériques Type Betacam Digital et **7516** DVD et **90** Disques Durs suivants les dispositions des cahiers des charges édité en 2011.

Les étapes de la numérisation comprennent le conditionnement des supports à transférer suivant un listing réalisé au préalable.

- 1- Examen des bandes : déterminer l'état de conservation, constat des symptômes de dégradation (humidité, collure, salissure, décomposition)
- 2- le nettoyage : déroulement et rembobinage, dépoussiérage avec des machines « Recortec » - (figure 6),
- 3- Restauration des collures mécaniques en cas de besoin.



Figure 6- Une machine de nettoyage Recortec pour bandes 2 pouces

Les bandes qui présentent fortes salissures et encrassement sont soumises au chauffage (baking) dans un four spécialisé en évitant le choc thermique suivant des normes précises pour augmenter la résistance des bandes à transférer (figure 7).



Figure 7- four pour les bandes 02 pouces

Les supports ainsi nettoyés sont admis à l'acquisition du signal vidéo sur des stations qui permettront la vérification de l'image et du son transféré : régulation du tracking, ajustement de la forme et du niveau du signal, vérification des pistes sonores et élimination des défauts d'enregistrements ou de lecture (drop-out).

A la fin de la procédure de transfert, la bande initiale est rembobinée, la finalisation du traitement des bandes 02 pouces consiste à visualiser les futurs montages, ces procédures sont assistés par des logiciels spécifiques, le logo de la Télévision Tunisienne est apposé sur l'ensemble des éléments vidéo.

Vient ensuite la phase de contrôle qualité pour l'examen de la qualité technique et vérification des procédures de transfert (durée, format, débit, aspect ratio, caractère mono/stéréo...).

Pour chaque fichier crée, un fichier de consultation est généré par le système de gestion des archives audiovisuelles ,c'est l'ultime étape : le visionnage documentaire qui consistera en une mise à jour des données soit par vérification, modification ou mention des métadonnées réelles enregistrées,(l'année d'enregistrement, le numéro, le titre, la durée, la date de numérisation, le nouveaux identifiants, mots clés, intervenants...)

Les listings préalables seront mis à jour sur des bases de données référentielles qui renferment les indices documentaires, techniques et artistiques des œuvres audiovisuelles.

Des centaines d'images d'archives continuent à être transmises depuis le mois de mai 2016 pour commémorer le cinquantième anniversaire de la télévision Tunisienne, 80 soirées d'archives ont été diffusées sur les deux chaînes nationales depuis 2015.

7-Œuvres télévisées : immortaliser le patrimoine audiovisuel

Rendre les œuvres télévisées pérennes implique qu'elles auront le statut du patrimoine national à toute œuvre diffusée par les canaux publics.

Les expériences de numérisation menées par la Télévision Tunisienne ont révélé que allocation des ressources financières, humaines et logistiques propres est une mission délicate, il est évident que la fragilité des supports reflète la fragilité du statut des œuvres télévisées du point de vue juridique.

Les fonds télévisuel est estimé à plus de 280.000 heures enregistrées sur les différents supports, les anciens supports susceptibles d'être numérisées représentent approximativement

la moitié du fonds actuel qui va progressivement évoluer vers l'obsolescence, la technologie chasse les fonds inactifs.

Il a été avancé que le corpus légal concernant ces œuvres est mince, il n'existe pas en Tunisie un groupement professionnel ni associatif pour représenter les professionnels de la gestion des œuvres audiovisuelles, les activités scientifiques concernant le domaine sont celles animées par l'institut supérieur de documentation de Tunis.

Pour la formation des professionnels de l'audiovisuel, l'institut supérieur assure des cursus spécialisés en archivage audiovisuel tout au long des cycles universitaires, l'institut des arts de multimédias forme des futurs praticiens de l'image animée, mesurant l'ampleur des projets de numérisation, il devrait consacrer une section des programmes à la gestion des anciens standards de diffusion et à leurs équipements, le savoir-faire en la matière vire à la disparition.

Seule la réglementation interne au niveau de la Télévision Tunisienne est présente, elle est le fruit des initiatives propres de l'établissement dans le cadre de protection et de sauvegarde du patrimoine télévisuel.

Le paysage audiovisuel en Tunisie est en progression active, la diffusion numérique concerne des dizaines de chaînes télévisées, de stations radiophoniques, la télévision tunisienne détient le monopole des droits de production et pourrait servir de source d'images, d'enregistrements pour ces nouvelles institutions nées durant le dernier quinquennat.

La Télévision Tunisienne a parrainé l'art cinématographique dès le début, elle fût la passerelle des professionnels du cinéma : réalisateurs, cinéastes, acteurs et autres spécialités de l'audiovisuel ont eu des expériences télévisuelles.

Une collaboration avec le centre national du cinéma et de l'image¹³³ créé en 2012, serait propice à plusieurs niveaux : échange d'équipements et d'expertise, activités scientifiques, les projets de numérisation auront plus d'impact et la coopération internationale offre des opportunités à plusieurs niveaux : aides financières, équipements, stages de perfectionnement, formations en restauration, recyclage pour les professionnels de l'audiovisuel.

¹³³(13) Décret no 2012-753 du 02 juillet 2012, fixant l'organisation administrative et financière et les modalités de fonctionnement du centre national du cinéma et de l'image, JORT n°55 du 13 juillet 2012.

Conclusion :

Les chaînes de télévision thématiques ont déjà tranché sur la valeur probatoire des œuvres télévisées et la périodicité des processus de numérisation, souvent privées elles ont misé sur les dits capitaux en s'enrichissant du savoir faire des experts des chaînes publiques et en cherchant l'appui des instances internationales de préservation du patrimoine audiovisuel.

Les organismes publics de l'audiovisuel ont la responsabilité de diffuser des œuvres à un public disparate et ensuite pérenniser l'accès à tout ceux qui réclament des références pour les besoins des productions télévisées, de recherche scientifique ,l'attestation des droits de diffusion et dans le cadre de partenariat avec d'autres organismes homologues.

Par constat, la numérisation des œuvres audiovisuelles est un processus semé de défis, les enjeux concernent la valeur culturelle, économique et patrimoniale des fonds à transférer.

Le processus de numérisation des œuvres télévisées est basé sur des expériences de transfert des plusieurs standards d'enregistrement qui existent à la Télévision Tunisienne exécutés par ses propres ressources ou grâce au concours de plusieurs instances spécialisées et programmes de coopération conçus pour la sauvegarde et la restauration du patrimoine audiovisuel mondial. Plusieurs paramètres concernent le processus de numérisation et sont actionnés par l'accroissement perpétuel de la volumétrie fonds des archives télévisées, les ressources humaines et matérielles suffisantes ne sont pas allouées systématiquement, l'obsolescence des supports suite à l'évolution technologique , la multitude des standards d'enregistrement (mécaniques, analogiques et numériques) ; imposent la mise en œuvre des mesures nécessaires et urgentes pour la pérennisation de l'accès et ainsi la diffusion des archives télévisées:

1- A l'échelle nationale et dans les meilleurs délais, attribuer aux archives de la Télévision Tunisienne le statut des **archives nationales de l'audiovisuel afin de préserver leur originalité et unicité** et créer ainsi de textes de lois spécifiques au secteur de l'audiovisuel, aux fonctions des œuvres télévisées,

2- **L'instauration d'une institution autonome chargée des stratégies de numérisation** des œuvres télévisées suivant la pratique du dépôt légal et à l'instar de l'institut national de l'audiovisuel en France (INA) semble être une solution à long terme pour les défis qu'impose la gestion des projets de numérisation des œuvres télévisées (restauration, transfert et mise à jour) .

La question remonte à Avril 2002 lors du séminaire de la Fédération Internationale des Archives Télévisées (FIAT) en Tunisie et demeure d'actualité ; la gestion des œuvres au quotidien (acquisition, description, catalogage et indexation) en vue de la diffusion au quotidien, prime sur la réflexion, la préparation, la sélection, l'exécution et le suivi des projets de numérisation des organismes publics de diffusion .

3- Il est essentiel pour la survie du fonds d'adopter une stratégie nationale et à long terme pour la numérisation des différents standards professionnels de diffusion au niveau de la Télévision Tunisienne et avec le concours des instances gouvernementales : le Ministère de la culture, les Archives Nationales de Tunisie (ANT) et le Centre National du Cinéma et de l'image (CNCI).

4-Renforcer la **formation des ressources humaines** (cadres techniques/cadres documentalistes) pour consolider le processus des archives télévisées, les institutions universitaires sont concernées pour la continuité des services fournies lors du transfert et de la mise à jour des données.

L'INA assure des cycles de formation aux métiers de l'audiovisuel et du multimédia pour les étudiants et les professionnels (Ina Sup, Ina Expert).

5-A l'échelle internationale, **fructifier les projets de coopération à l'instar de capmed et Medmem** dans le cadre **d'Euromed heritage**, avec le concours de l'INA les projets de l'UNESCO de la FIAT et autres groupements professionnels des archivistes de l'audiovisuel des pays qui ont acquis une maturité en matière de gouvernance des politiques et stratégies de numérisation, l'Ina a élaboré sa stratégie à long terme de 1999 à 2015.

Bibliographie :

(1) les œuvres citées sont les émissions phares de la télévision tunisienne,

(2) décret 2007-1868 du 27 juillet 2007 portant création, organisation administrative et financière et modalités de fonctionnement de la Télévision Tunisienne, Journal Officiel de la République Tunisienne (JORT) n°60, pp2602-2610.

(3) Helen P. Harrison ; Les archives audiovisuelles à travers le monde, document en ligne pp194-201. (Consulté le 27 mai 2016).

<http://www.unesco.org/webworld/wirerpt/wirfrench/chap14.pdf>

(4) Marie-Anne Chabin, qu'est ce qu'une archive audiovisuelle ? in E-dossier de l'audiovisuel : L'Extension des usages de l'archive audiovisuelle, pp1-5.

(5) Wahid Gdoura القوى العاملة في مجال الأرشيف السمعي البصري: الكفاءات والمؤهلات in, revue ASBU, n°4, 2010, pp58-69.

(6) Mohamed Almanasir, الحفظ والصيانة لأرشيفات الصوت و الصورة, in revue ASBU, n°4, 2010, pp37-57.

(7) Dominique Saintville, la stratégie de sauvegarde et de numérisation des archives de l'Institut National de l'Audiovisuel, 1995-2015, in preservation news, n°47, Mai 2009, pp18.

(8) le projet CapMed est issue d'une convention signée entre l'Etablissement de la radiodiffusion télévision Tunisienne (ERTT) et l'institut national de l'audiovisuel (INA) France.

www.ina.fr – www.euromedaudiovisuel.net

(9) la convention entre la Télévision Tunisienne (TT) et l'Institut national de l'audiovisuel (INA) fut signée en Mars 2011. <http://www.medmem.eu/>

(10) lexique marc <http://www.marc21.ca/MaJ/BIB/B007-EnregistrementVideo.pdf> consulté le 28 Mai 2016.

(11) Décret no 2012-753 du 02/07/2012, fixant l'organisation administrative et financière et les modalités de fonctionnement du centre national du cinéma et de l'image, JORT n°55 du 13 juillet 2012, pp 1632-1639.

© Actes de la première édition du colloque international sur les bibliothèques et archives à l'ère des humanités numériques : Archivage numérique et réseaux.

Ouvrage édité et distribué par ISD

Impression : ImprimArt
ISBN : 978-9973-904-13-3