HAL
open science

Corpus des Archives à l'Ere des Humanités Numériques : Etude d'Evaluation.

Ben Fatma, Abderrazak Mkadmi

To cite this version:

Ben Fatma, Abderrazak Mkadmi. Corpus des Archives à l'Ere des Humanités Numériques : Etude d'Evaluation.. Revue des Sciences Humaines et Sociales, 2018. hal-02997839

HAL Id: hal-02997839

https://hal.science/hal-02997839v1

Submitted on 10 Nov 2020

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

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

Corpus des Archives à l'Ere des Humanités Numériques : Etude d'Evaluation.

مجموعات الأرشفة في عصر العلوم الإنسانية الرقمية

دراسة تقييمية

Fatma Ben Amor; fatma_bilel@yahoo.fr

Abderrazak Mkadmi; amkadmi@gmail.com

Unité de recherche BNP, ISD, **Université de la Manouba**

تاريخ استلام البحث: 26-05-2018 **تاريخ قبول البحث: 22-06-2018**

ملخص:

مع التطور المذهل لتكنولوجيا المعلومات والاتصالات على شبكة الإنترنت، شهدت العلوم الإنسانية والاجتماعية تحولات جذرية في أساليبها ومرتكزاتها، حيث أصبحنا نتحدث عن مفهوم جديد وهو "الإنسانيات الرقمية" وقد أدخل تغييرات كبيرة في جميع المجالات وخاصة في مجال المكتبات والتوثيق والأرشيف.

وفي هذا السياق، تتطرق مقالتنا إلى موضوع العلوم الإنسانية الرقمية وتأثيرها على الأرشيف والأرشيفيين، مع التركيز على دراسة منتجات الإنسانية الرقمية، وهي أرصدة الأرشيف الرقمية، وأدواتها ووظائفها وطرق حفظها على المدى الطويل. حيث نقدم منهجية دراستنا والنتائج الرئيسية المتعلقة على وجه الخصوص بتأثير الانسانيات الرقمية على تطور مفهوم الأرشيف وممارسات معالجته والحفاظ عليه على المدى الطويل.

Abstract

With the spectacular development of web-based information and communication technologies, Human and Social Sciences (SHS) have experienced a profound upheaval in their methods and paradigms. We are talking about a new concept which is "the digital humanities" that has introduced major changes in all areas and especially in the area of libraries, documentation and archival science.

It is in this sense that our article deals with the issue of digital humanities and its impact on archives and archivists with an emphasis on the study of digital humanities products, namely the corpuses of digital archives, their archival tools and features and their long-term preservation methods. We present the methodology of our study and the main results related in particular to the evolution of the notion of archives, the practices of their treatment and their long-term conservation.

Introduction

Les sciences humaines et sociales (SHS) sont un champ disciplinaire qui traite de l'homme et de ses relations avec son environnement naturel, culturel, social, politique et économique. Elles « réunissent un ensemble de disciplines scientifiques hétérogènes qui traitent de l'humain, des organisations, des échanges » (Djamel, 2013, p.1). Elles couvrent les sciences sociales caractérisées par une activité de chercheurs menée en groupe et se rapprochant des sciences expérimentales et des sciences humaines (lettres, langues, littératures, philosophies, etc.) qui restent dans leur modèle de tradition classique et d'activité basée sur les sources primaires et le travail individuel (Djamel, 2013, p.1)

Contrairement aux sciences exactes, les SHS étaient considérées comme des disciplines qui " traitent de l'humain, des organisations, des échanges, de la communication" (Djamel, 2013, p.1). Elles sont caractérisées par leur complexité liée à la structure et à la nature des êtres humains. Cette complexité rend difficile l'association de l'humaniste au numérique « Est-il possible qu'une chose soit à la fois numérique et humaniste ? » (Scholes, 2008). En effet, 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 parait que cette idée est fausse étant donné que, parmi tous les arts la littérature a paradoxalement rencontré l'informatique le plus tôt (Clément, 2001, p1).

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 touchant leurs méthodes et leurs paradigmes.

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

Avec les humanités numériques, des changements majeurs ont touché non seulement les disciplines concernées à l'instar de l'histoire,

de la littérature, de la linguistiques, de la sociologie mais ont également influencé les secteurs concernés comme la visualisation, l'édition électronique, les bibliothèques, la documentation et l'archivistique.

Pour les archives, les technologies numériques introduites dans les humanités ont permis de développer de nouveaux outils et applications qui favorisent l'accès aux archives et leur mise en valeur influençant leurs valeurs, leur fonctionnement et leurs rôles ainsi que ceux des archivistes.

C'est dans cette optique que s'inscrit notre article qui traite la question des humanités numériques et ses impacts sur les archives et les archivistes en mettant l'accent sur l'étude de produits des humanités numériques à savoir les corpus des archives numériques.

I- Problématique

Les humanités numériques constituent un champ d'expérimentation qui fait appel d'une part à de nouveaux rapports entre technologies numériques et sciences humaines et sociales, et qui relie d'autre part des professions distinctes et diverses disciplines à savoir celles liées à la documentation, aux bibliothèques et aux archives.

Les archives et le métier d'archiviste ont subi des transformations face au développement des technologies de l'information et de la communication et face à l'évolution numérique constante. Avec l'avènement des humanités numériques, les archives se trouvent engagées dans la recherche en sciences humaines confrontée à de nouveaux moyens de recherches et de nouvelles applications.

Certes, les bibliothèques ont joué un rôle important dans la mise en place et l'application de projets relevant des champs 'humanités numériques', à savoir :

- La participation à la création ou à l'amélioration des notices Wikipédia (ex : British Library) ;
- La numérisation et la mise en ligne de sources (les actions de la b.n.f. : Gallica).

Ce qui montre que les humanités numériques sont à la fois une opportunité aussi bien pour les chercheurs que pour les bibliothécaires, mais est-ce vraiment le cas pour les archives ?

Les archives sont, sans aucun doute, l'objet de changements importants à l'ère des humanités numériques. Elles doivent faire face à de nombreux défis et doivent réévaluer leur rôle en tant que gardiens du patrimoine culturel. Contrairement aux bibliothèques, dont l'accès est au premier plan, les archives, étaient considérées beaucoup plus comme les gardiens de collections soigneusement forgées.

Dès lors, et vu ce rôle de gardiens, les archives se trouvent confrontées à un phénomène où une quantité importante du patrimoine culturel est perdue en raison de la numérisation rapide de l'information et le manque de méthodes fiables de conservation à long terme ; ce qui pousse les archivistes et les experts en informatique à exprimer leurs préoccupations sur « l'âge sombre » dit numérique. (Viiri, 2014, p.6)

Ainsi se pose la question sur la mise au point des procédures, des critères et des infrastructures de sélection et de préservation à long terme des documents numériques « natifs » qui sont en évolution constante.

En outre, il faut mettre l'accent sur les nouveaux rôles que doivent assumer les archives quant à la constitution, à l'analyse et à l'organisation des corpus numériques.

Ensuite, c'est la question de l'accès libre à la mémoire culturelle qui doit être ouvert tant que possible au public [1] , or du fait de leur statut public et privé les archives posent un certain nombre de difficultés juridiques liées au respect des lois en vigueur (les droits des personnes, des créateurs, des propriétaires et des utilisateurs).

A ce titre, Adelheid Heftberger, dans un travail sur les archives et les humanités numériques, pense qu'il faut plutôt mettre l'accent sur le rôle des archives et des archivistes dans le développement des projets humanités numériques et que la numérisation et l'édition électronique fournissent de nouveaux domaines de l'archivage (Adelheid, 2014). Quant à Presner et Schnapp, dans « The Digital Humanities Manifesto 2.0 » ils ont considéré l'archive comme un partenaire beaucoup plus proactif qui doit devenir un lieu d'enseignement et d'apprentissage « *An*

[1] Article 9 du manifeste des *digital humanities*

archive must become a place of teaching and hands on learning » (Schnapp, 2009)

Pour sa part, Sampo Viiri affirme dans un rapport publié à Londres, par l'institut finlandais sur les sciences humaines numériques et la préservation du patrimoine culturel numérique, considère que le rôle de l'archiviste est critique dans la démarche de recherche de l'histoire comme un moyen de chemin d'accès au contenu (Viiri, 2014, p.4).

C'est dans ce sens qu'il y a lieu de se poser la question suivante :

Dans quelle mesure les humanités numériques font-elles évoluer les archives, leurs outils et fonctionnalités archivistiques, et leurs méthodes de conservation à long terme ?

Pour répondre à cette problématique, nous avons suivi la méthode de collecte des données par la grille d'évaluation des corpus d'archives numériques.

II- Méthodologie

Dans le cadre de notre recherche visant à étudier l'impact des humanités numériques avec leurs méthodes, leurs techniques et leurs outils dans les champs de la gestion des archives numériques et leur conservation à long terme, nous avons mené une étude évaluative sur 20 corpus d'archives numériques, et ce à l'aide d'une grille d'évaluation pour juger le rôle des techniques et logiciels orientés vers les humanités numériques dans la valorisation des archives. Cette étude a pour objectif principal de voir comment, avec la mise en ligne des archives, les humanités numériques, ont changé les méthodes de traitement, de description, de conservation et de diffusion des archives numériques.

1. La démarche de l'élaboration de la grille :

Nous avons essayé d'établir une grille d'évaluation des corpus d'archives numériques scientifiques applicable aux spécificités des archives numériques. Cette grille nous permet d'analyser d'une manière systématique et objective les caractéristiques qui distinguent les corpus d'archives numériques en tant que site web par rapport à d'autres sites, ainsi que celles qui les distinguent en tant qu'archives (au niveau des fonctionnalités, conservation pérenne, nouvelles méthodes de

description) et dresser les changements subis par les archives à l'ère des humanités numériques par l'évaluation de leur mise en réseaux.

L'élaboration de la grille d'évaluation nécessite plusieurs étapes :

- Identification des critères d'évaluations possibles.
- Consolidation des critères.
- Catégorisation des critères.
- Etablissement d'un test de la grille d'évaluation.

1.1. L'identification des critères d'évaluation :

Pour arriver à cerner une liste idéale de critères d'évaluation, il était important d'effectuer un inventaire de critères nécessaires utilisés pour l'évaluation. Dans ce sens, il était utile de consulter plusieurs sources pour identifier les critères d'évaluations.

Parmi les sources les plus importantes au niveau de l'évaluation pour la bonne qualité web, nous avons consulté :

- Le site de OPQUAST[2] (Open Quality Standards) : c'est une base de données qui englobe une liste de check-lists en tant qu'outils d'évaluation de sites web. OPQUAST élabore et diffuse des référentiels publics de bonnes pratiques pour la qualité Web[3].

La dernière liste présentée par OPQUAST est mise à jour en 2015 (la première version en 2010, la deuxième en 2004) Dates à vérifier s'il vous plaît, il doit y avoir une confusion.

Après la consultation, nous avons fait un travail de sélection, d'ajouts et de suppressions de certains critères. Nous avons surtout fini par valider ce choix à l'aide de la liste complète présentée par OPQUAST.

1.2. La consolidation des critères d'évaluation

A cette étape, nous avons essayé d'uniformiser les critères regroupés, d'éliminer les critères en double et d'unifier la dénomination des critères ayant la même fonction. Les autres critères inutilisables ou

[2] http://checklists.opquast.com/fr/

insignifiants ont été éliminés tels que les publicités ou pop-up qui nuisent à la navigation et qui n'apparaissent jamais sur les sites d'archives numériques. Cette procédure a permis de réduire la liste originale à 115 indicateurs de qualité.

1.3. La catégorisation des critères d'évaluation

Après l'uniformité des critères, nous avons procédé à l'organisation et à la hiérarchisation des critères ayant le même angle (ex : sous la présentation visuelle, nous avons inséré la visualisation de corpus, la charte graphique et la visibilité textuelle). Puis nous avons procédé à la codification des critères. Tous ces critères ont été formulés par des questions fermées formulées ? de telle sorte que nous puissions cocher la réponse affirmative.

2. La démarche de choix de corpus d'archives numériques

Nous sommes arrivés à établir une grille de 20 sites de corpus d'archives numériques. Nous avons essayé d'obtenir un équilibre relatif entre la typologie de ces corpus d'archives (nationales, départementales, privées, ou autres…) et le nombre de sites utilisant tel ou tel logiciel par rapport à d'autres.

Nous avons choisi d'abord les corpus d'archives rassemblées par les institutions publiques qui ont pour mandat la gestion des archives définitives pour dresser le traitement numérique des sources anciennes par les moyens et les outils des humanités numériques. Dans ce sens, nous avons choisi trois archives nationales importantes dans le monde qui sont :

- Les Archives Nationales de France (ANF) (http://www.archives-nationales.culture.gouv.fr/)
- Les Archives Nationales des Etats Unis : National Archives and Record Administration (NARA)(https://www.archives.gov/)
- Les Archives Nationales d'Australie : National Archives of Australia (NAA)
- (http://www.naa.gov.au/)

Nous avons choisi aussi d'autres corpus d'archives numériques qui sont présentées par des sites des archives départementales à l'instar des :

- Archives départementales de Gironde (http://archives.gironde.fr/)
- Archives départementales du Vaucluse (http://archives.vaucluse.fr/accueil/)

- Archives départementales Mission Basse Bretagne (http://bassebretagne-mnatp1939.com/)
- Archives départementales Val de Marne (http://archives.valdemarne.fr/content/archives-en-ligne)
- Archives départementales Ardennes (http://archives.cd08.fr/article.php?larub=3)
- Archives départementales de l'Aube (http://www.archives-aube.fr/)
- Archives départementales Eure et Loire (http://www.archives28.fr/article.php?larub=1)
- Archives départementales Yvelines (http://archives.yvelines.fr/)
- Archives départementales Morbihan (http://www.archives.morbihan.fr/accueil/)
- Archives départementales Aisne (http://archives.aisne.fr/)
- Archives départementales Puy de-dôme (http://www.archivesdepartementales.puydedome.fr/)
- Archives départementales Loire Atlantique (https://archives.loire-atlantique.fr/jcms/chercher/archives-numerisees-fr-c_5562)

Ainsi, et dans le but de diversifier la typologie des archives numériques, nous avons choisi d'autres corpus d'archives numériques privées conservées par des structures publiques ou privées à l'exemple de :

- Henri Poincaré. Laboratoire d'Histoire des Sciences et de Philosophie (http://henri-poincare.ahp-numerique.fr/)
- Archives Bourbaki : source de l'histoire (http://archives-bourbaki.ahp-numerique.fr/
- Archives Christiane Taubira (http://taubira.anaphore.org/pages/infos.html)
- Archives numérique Disanti (http://archive.desanti.huma-num.fr/)
- Les carnets géologiques de Philippes Glangeaud (http://bibliotheque.clermont-universite.fr/glangeaud/)

Pour extraire cette liste des archives départementales, nous avons recouru au portail des archives départementales en ligne[4] qui nous a facilité l'accès à ces archives en ligne.

Ainsi, nous avons choisi les sites dont la création est basée sur des plateformes d'archivage numérique (PLEADE, OMEKA, ARKOTHEQUE..) parce qu'elles assurent la préservation numérique, le référencement sur le long terme des documents numériques et l'interopérabilité (OAI-PMH, OAI-ORE)[5]

[4] http://www.archives-departementales.com/

[5] https://archivengines.wordpress.com/2012/05/02/difference-plateforme-archivage-numerique-gestionnaire-contenu/

Ainsi nous avons concentré notre choix de critères de corpus d'archives numériques sur le ou les outils utilisés (outil de travail collaboratif, …).

L'élaboration d'une grille d'évaluation est un outil nécessaire qui va nous permettre de collecter les informations nécessaires sur les 20 corpus d'archives numériques, de recueillir ces informations, de les schématiser sous forme de courbes et histogrammes pour en analyser les résultats.

III- Les Résultats de l'évaluation

Nous présentons dans cette partie les résultats des données collectées par la grille d'évaluation et les analyses de ces résultats au niveau de l'interface et du contenu pour traiter les tendances ergonomiques et fonctionnelles des corpus d'archives numériques sollicités.

1. L'évaluation de l'interface :

Dans l'évaluation de l'interface des corpus numériques, nous avons abordé deux thématiques : la présentation visuelle et la navigation.

Pour la présentation visuelle, les résultats de l'évaluation montrent que la majorité des corpus d'archives numériques sont caractérisés par une presentation visuelle claire et harmonieuse jouant un rôle important dans l'affichage organisé d'informations déjà structurées et dans la perception visuelle de l'information ; ce qui rejoint l'idée de Johanna Drucker (Drucker, 2004) qu'il y a une conscience de l'importance du design graphique et visuel dans les projets des humanités numériques et que le design graphique n'est pas un simple exercice accessoire.

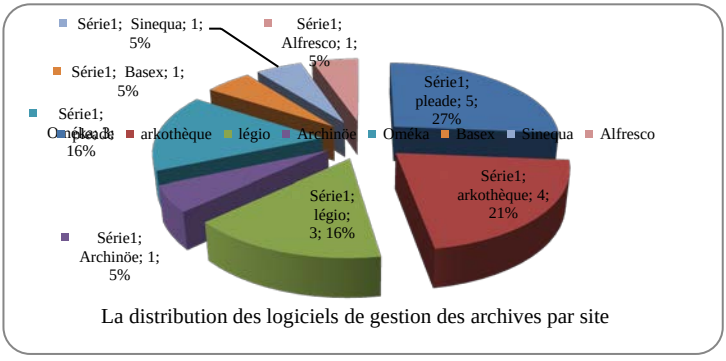
Tableau 1 :
l'identité de corpus des archives numériques

Identité du corpus	Nombre de sites	%
Le corpus est-il créé par une plateforme informatique	19	95%
Le corpus comporte-t-il un logo ?	19	95%
Si oui, le logo est-il facile à mémoriser ?	16	80%
L'extension utilisée dans l'adresse url est-elle cohérente avec l'identité, l'activité, la zone géographique couverte ou avec le nom du domaine ?	20	100%
Le graphisme choisi est-il cohérent avec l'image de la marque ?	20	100%

Selon le tableau n°1, la totalité des corpus d'archives (100%) présentent une identité claire et cohérente que ce soit au niveau de leur adresse url qui traduit leurs activités, ou les zones géographiques couvertes et le nom du domaine ou le graphisme choisi qui reflète leur image de marque. 95% de ces corpus d'archives comporte un logo dont la majorité est facile à mémoriser. Ainsi 95% des corpus d'archives sont créés par des plateformes de gestion archivistique.

La distribution des logiciels utilisés (figure 1) montre l'importance donnée par ces corpus à des plateformes orientées gestion d'archives telle la plateforme pleade utilisée par 27% des corpus d'archives numériques surtout départementales et en deuxième lieu le logiciel Arkothèque dont 21% utilisent des corpus d'archives numériques.

Figure 1 :
La distribution des logiciels de gestion d'archives

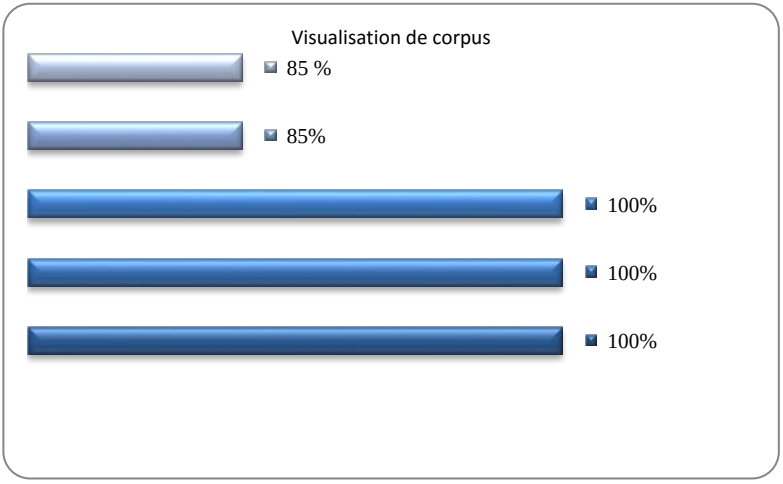


La distribution des logiciels de gestion des archives par site

Vu l'importance du travail collaboratif et de partage de l'information, les services d'archives numériques se trouvent obligés de choisir des plateformes qui offrent des fonctionnalités collaboratives à l'instar de pleade, Arkothèque et Ligéo.

Pour la visualisation, la totalité des corpus d'archives numériques sont clairs, lisibles et présentent une homogénéité au niveau de toutes les pages et au niveau du choix et de l'organisation des couleurs (figure 2). Cette importance donnée à la visualisation générale du corpus est justifiée par la stratégie suivie par les services d'archives numériques pour attirer l'attention des utilisateurs et pour valoriser les fonds numériques accessibles en ligne.

Figure 2 :
La visualisation de corpus d'archives numériques



Si les taux de corpus d'archives numériques sont d'autant plus marqués au niveau de la visualisation, c'est qu'ils respectent une charte graphique dans leurs sites. En effet, 100% des corpus d'archives numériques présentent une charte graphique cohérente, parmi eux 95% utilisent des éléments graphiques qui guident la visite (flèche), et qui ajoutent de la valeur au contenu (tableau n°3).

Tableau 2 :
La charte graphique

Charte graphique	Nombre de site	%
La charte graphique est-elle cohérente sur l'ensemble des sites ?	20	100
Les éléments graphiques guident-ils la visite ?	19	95
Les éléments graphiques ajoutent-ils de la valeur au contenu ?	19	95
Les textes, images et documents multimédia sont-ils correctement agencés ?	19	95
Le corpus utilise-il des images symboliques ou moins représentatives ?	12	60

Pour la navigation, la plupart des corpus d'archives numériques (entre 95% et 100%) présentent une navigation cohérente et intuitive grâce à une bonne structuration qui fournit à l'utilisateur une représentation mentale adéquate (Tableau n°4). Ces taux sont d'autant plus marqués avec les corpus d'archives qui utilisent le plus d'outils de navigation.

Tableau 3 :

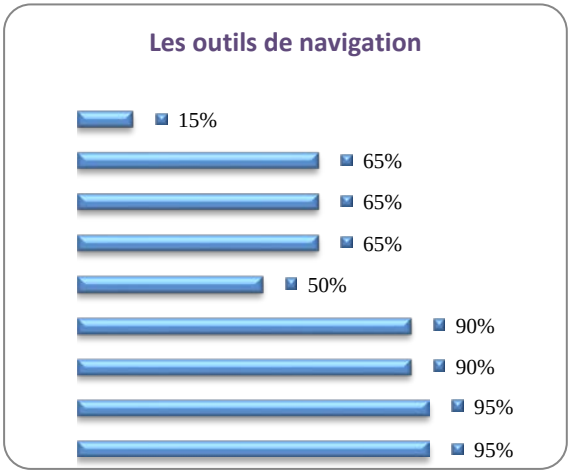
la structuration des corpus d'archives numériques

Structuration	Nombre de site	%
L'adresse du site est-elle apparente et courte ?	17	85
Le chargement de site et des pages est-il rapide ?	20	100
Y'a-t-il un plan de site ?	15	75
Peut-on accéder facilement au plan du site ?	15	75
Peut-on accéder au plan du site depuis chaque page ?	15	75
La navigation dans le site se fait-elle facilement (aller-retour, bouton "accueil" présent…) ?	20	100
L'information est-elle facile à retrouver ?	19	95
Les pages sont-elles clairement organisées ? (différentes parties distinctes…)	20	100
Savez-vous toujours où êtes-vous ? (présence de repères…)	19	95

En effet, la majorité des corpus d'archives (entre 90 et 95%) présentent une mise en page claire qui facilite la navigation à l'aide de l'utilisation de moteurs de recherche efficaces et d'outils d'aide à la navigation (boutons compréhensibles, file d'Ariane, des flèches d'orientation...). 65% présentent des liens validés et actifs, qui pointent vers des sites opérationnels même si 15% seulement des sites

d'archives respectent une présentation qui utilise les standards de présentation de ces liens (Des standards qui font partie des normes d'accessibilité et qui consistent, entre autre, à attribuer une couleur différente aux liens visités et non-visités, différencier les liens et les rendre facilement identifiables et donner un libellé explicite à tous les liens).

Figure 3 :
Outils de navigation dans les corpus d'archives numériques



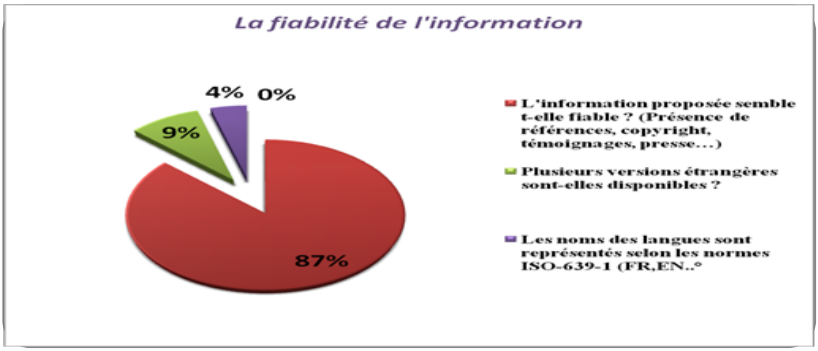
2. L'évaluation des corpus (contenu)

Pour évaluer le contenu des corpus, nous avons utilisé les critères de la crédibilité, de l'interactivité et des fonctionnalités archivistiques.

Pour la crédibilité, les sites de corpus d'archives numériques sont des sites informatifs, leurs objectifs consistent à présenter les ressources, le patrimoine mobilier et les collections du département ou de l'institution concernés. Les informations contenues sur le site sont issues de sources considérées comme fiables. C'est dans ce sens que 87% des corpus d'archives numériques sont qualifiés fiables au niveau de l'information proposée (présence de références, copyright, témoignages ...) (figure 4).

Figure 4 :

la fiabilité de l'information dans les corpus d'archives numériques



Pour la responsabilité intellectuelle**,** 100% des corpus mentionnent les noms, les logos et les références des responsables indiquant les sources des documents (tableau n°5). Certains corpus publient les images sous licence libre à l'exemple des carnets géologiques de Philippe Glangeaud.[6]

Tableau 4 :

la responsabilité intellectuelle

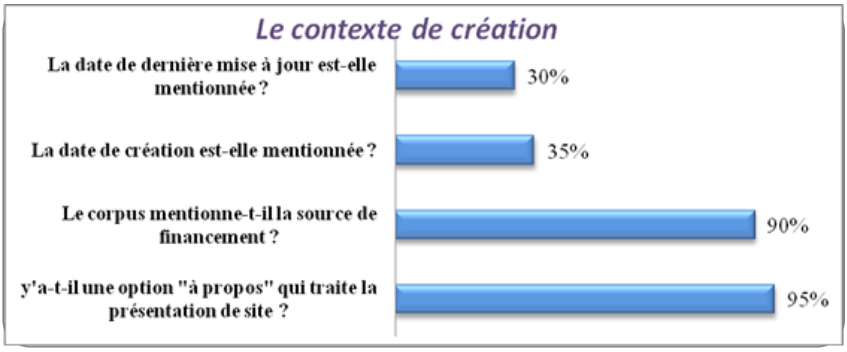
Responsabilité intellectuelle	Nombre de site	%
Les noms, logos et références des responsables sont-ils mentionnés ?	20	100%
Les sources des documents sont-elles indiquées ?	20	100%
Les images sont-elles accompagnées d'un copyright ?	17	90%

Pour **le contexte de création** des corpus d'archives (date de création et de mise à jour, sources de financement), la majorité des corpus d'archives numériques (entre 90% et 95%) mentionnent la source de leur financement. 35% mentionnent leur date de création et 30% mentionnent leur date de mise à jour (figure n°5).

[6]Philippe Glangeaud (né en 1866, mort en 1930) est un géologue français. Il fut professeur de géologie à l'Université de Clermont-Ferrand, collaborateur principal du Service Géologique de France et correspondant de l'Académie d'Agriculture de France.

Figure 5 :

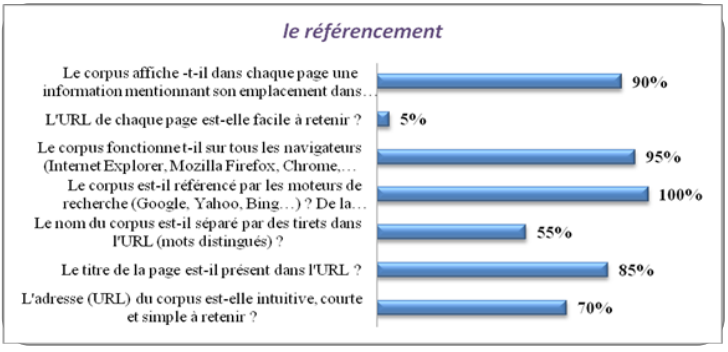
le contexte de création des corpus d'archives numériques



Pour le référencement, la totalité des corpus d'archives numériques évaluées ont été référencés de la même façon par les moteurs de recherche (Google, Yahoo, Bing…) et fonctionnent presque sur tous les navigateurs (Internet Explorer, Mozilla Firefox, Chrome, Opera, Safari). Entre 70 et 90% des corpus utilisent une adresse URL intuitive et courte dont le titre de la page est présent dans l'URL (figure n°6).

Figure 6 :

évaluation de référencement des corpus d'archives numériques



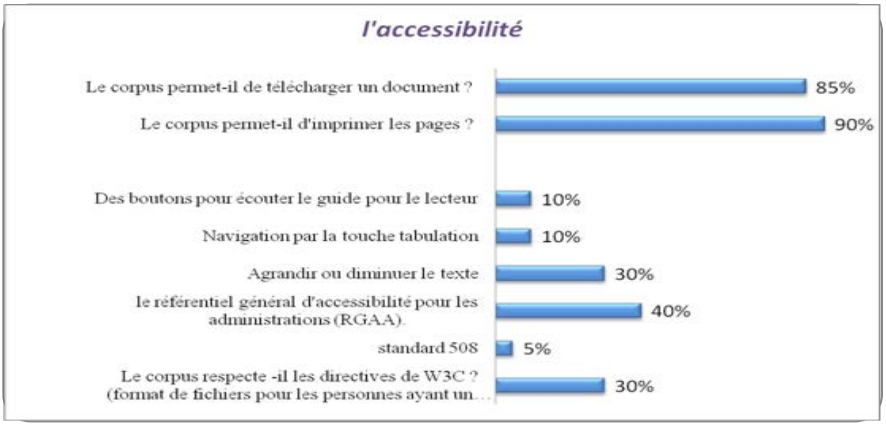
Pour **l'accessibilité**[7] , la plupart des sites d'archives suivent les préconisations faites par le W3C et les recommandations

[7] L'accessibilité exige que les projets et les savoirs numériques devraient être disponibles pour autant de personnes que possible en prenant en considérations les personnes handicapées (moteurs, cognitifs, sensoriels), les terminaux d'accès

d'accessibilité. 5% seulement des corpus équivalent à un seul corpus d'archives celui des Etats Unis (NARA) qui applique le standard 508[8] spécifique aux Etats Unis. En contrepartie, 40% des corpus appliquent le référentiel général d'accessibilité pour l'administration (RGAA)[9] dont 30% de ces corpus utilisent l'outil d'agrandir et de diminuer le caractère du texte et 10% des corpus utilisent d'autres outils tels l'utilisation des boutons pour écouter le texte ou la navigation par la touche tabulation (figure n°7).

Figure 7 :

évaluation de l'accessibilité des corpus d'archives numériques



Outre l'accessibilité, l'accès à l'information contenue dans ses fonds est l'une des grandes missions d'un service d'archives. En effet l'accès et l'exploitabilité des archives montrent leur utilité à la société, c'est dans ce sens que les services d'archives fixent des règles et présentent des conditions d'accès à leurs corpus d'archives.

(compatibilité avec des différents terminaux : micro-ordinateurs PC ; Mac ; Smartphones..) et les débits de connexions (la facilité d'affichage du haut débit (DSL).

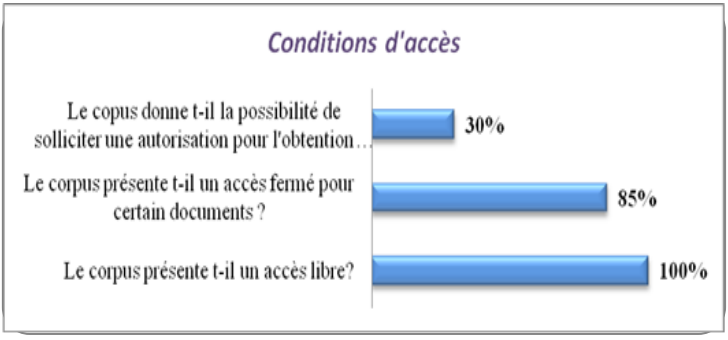
[8] La section 508 c'est la réglementation qui gère l'accessibilité aux Etats Unis, c'est une loi fédérale exigeant que toutes les technologies électroniques et informatiques développées, achetées, maintenues ou utilisées par le gouvernement fédéral soient accessibles aux personnes handicapées.

[9] Le Référentiel général d'accessibilité pour les administrations a pour objectif de qualifier l'accessibilité des contenus web et des applications présentés sous forme numérique et consultables à travers un navigateur web, quel que soit le support.

Selon notre évaluation, la totalité des corpus des archives numériques présentent un accès libre aux archives autorisant la réutilisation des informations contenues sur leur corpus à condition de ne pas faire un usage commercial des données, de ne pas diffuser publiquement ou à des tiers à l'exception d'un usage privé ou interne. 85% des sites d'archives présentent un accès fermé pour les archives privées (figure n°8).

Figure 8 :

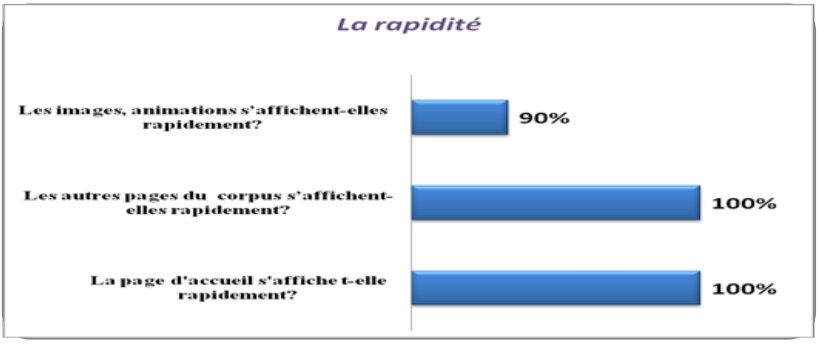
les conditions d'accès aux corpus d'archives numériques



Les corpus d'archives numériques sont aussi qualifiés performants au niveau de l'interactivité avec les visiteurs. En effet, 100% des corpus d'archives présentent une rapidité importante au niveau de l'affichage de la page d'accueil et les autres pages de corpus. Ainsi la majorité de ces corpus affichent leurs images et animations rapidement (figure n°9).

Figure 9 :

la rapidité des corpus d'archives numériques



En outre, 100% des corpus d'archives numériques mettent à la disposition des visiteurs des moyens de contact avec le responsable (coordonnées postales, coordonnées téléphoniques, email) (figure n° 10).

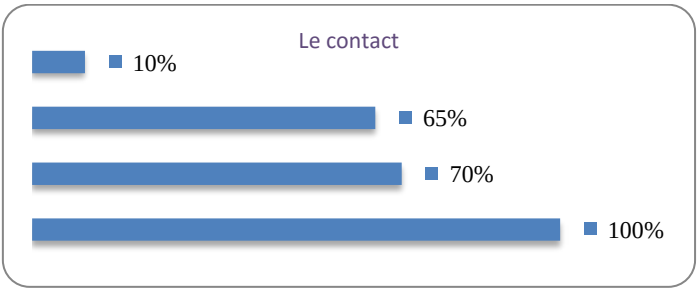


Figure 10 :
le contact dans des corpus d'archives numériques

Pour les fonctionnalités archivistiques des corpus d'archives numériques évalués, elles présentent une évolution remarquable au niveau :

- de la description par les métadonnées.
- de la recherche et la présentation des résultats.
- des outils de conservation.
- des outils de lecture développés
- du travail collaboratif et de partage.

La description des corpus d'archives numériques est marquée par l'utilisation importante des pratiques standardisées. Parmi les standards les plus utilisés par les corpus d'archives numériques l'EAD . L'EAD est appliqué par 76% des corpus d'archives numériques (dont la plupart sont en France), les métadonnées EAD sont jugées compatibles avec les pratiques françaises et le standard de description international ISAD (G). Elles représentent un bon outil pour les répertoires et les inventaires.

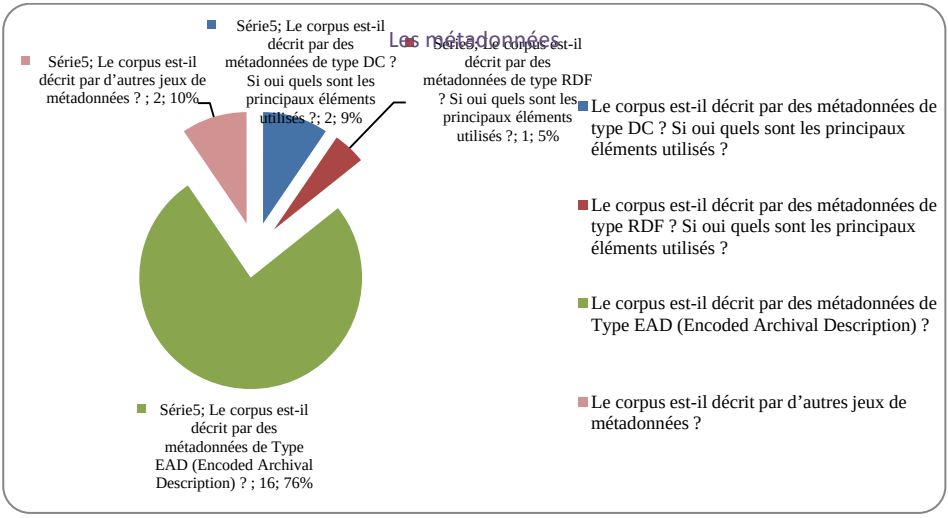
9% des corpus sont décrits par DC[10]. C'est le cas des archives de Mission Basse-Bretagne 1939 qui utilisent les 15 éléments de base de

[10] Dublin Core

DC. D'autres corpus d'archives numériques (10%) sont décrits par d'autres jeux de métadonnées (figure n° 11).

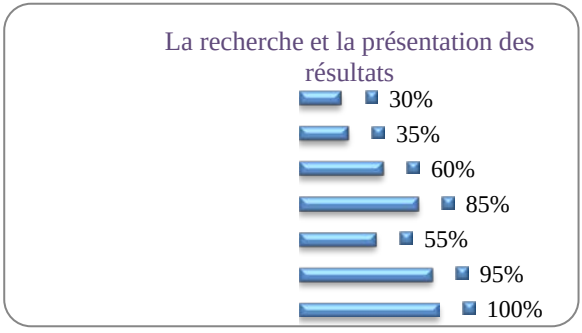
Figure 11 :

les métadonnées utilisées dans les corpus d'archives numériques



Pour l'évaluation des outils de recherche des corpus d'archives numériques, notre étude montre que la totalité des corpus d'archives numériques (100%) utilisent la recherche par mots clés, la majorité (entre 85% et 95 %) utilisent la recherche par collection spécialisée et la recherche avancée et entre 30 et 60 % des corpus d'archives utilisent les opérations booléennes et la recherche thématique parallèlement à la géolocalisation des résultats de recherche (figure n°12).

Figure 12 :
la recherche et la présentation des résultats

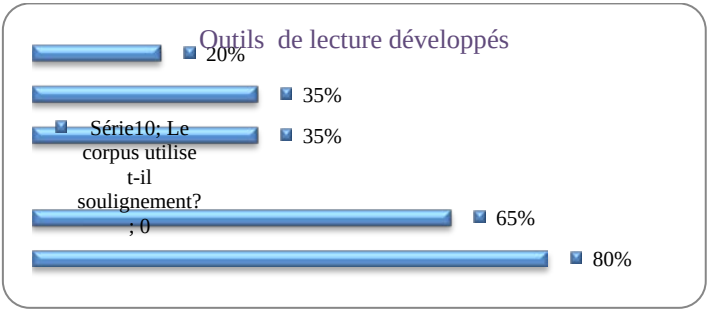


Les corpus d'archives numériques mettent à la disposition des lecteurs de nouveaux outils développés qui facilitent la lecture et l'exploitation des résultats de recherche lors de la consultation. L'étude montre que 80% des corpus d'archives utilisent l'outil de visionnement des documents numérisés. Cet outil est édifié par l'utilisation d'une interface développée de consultation des documents (figure n°13).

65% des corpus utilisent le surlignage lors de la recherche, cette technique permet de surligner les mots clés en affichant les résultats de recherche. 20% permettent de laisser des commentaires sur le contenu proposé. Pour l'annotation et la transcription, 30% des corpus l'utilisent en tant qu'outil de lecture numérique mais aussi en tant qu'activité collaborative que nous allons analyser dans la section suivante.

Figure 13 :

Les outils de lecture développés



Pour le travail collaboratif et le partage sur le réseau, la majorité des corpus d'archives numériques sont signalés dans les réseaux sociaux et proposent des fonctions collaboratives et de partage telles l'utilisation de la folksonomie, l'annotation collaborative, et même la transcription collaborative ainsi que d'autres activités qui facilitent le travail archivistique fastidieux avec la collaboration des archivistes et même avec des internautes intéressés par les archives. 65% des corpus d'archives numériques sont signalés sur des réseaux sociaux (Facebook et twitter...) sur lesquels ils partagent ou permettent aux lecteurs de partager le contenu des notices, et les résultats de recherche ou les images conservées. 5% de ces corpus valent un seul site celui de National Archives and Record Administration (NARA) donnent la possibilité de partage sur 189 réseaux sociaux et professionnels ce qui justifie leur impact sur la communauté sociale et professionnelle et

donne une image résolument moderne des activités des archives numériques tournées vers le « web 2.0 » (figure n°14).

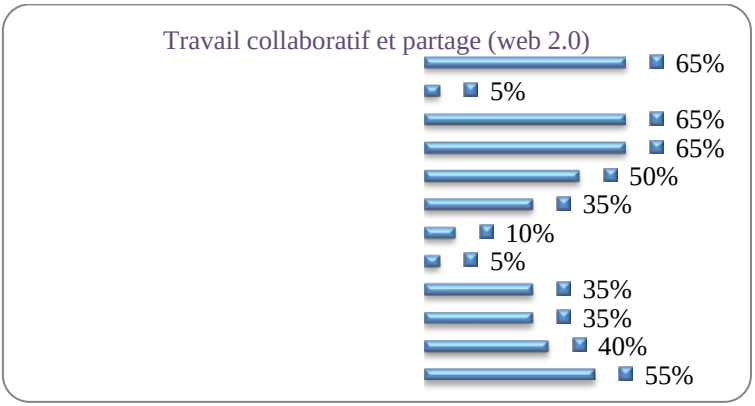


Figure 14 :

le travail collaboratif et partage dans les corpus d'archives numériques

Outre l'utilisation des flux RSS (35% des corpus), newsletters (50%), certains corpus d'archives recourent à la création d'un espace personnel à leurs utilisateurs, avec lequel les internautes peuvent participer aux activités collaboratives à l'égard de la transcription, les annotations et l'indexation collaborative. 35% des corpus d'archives numériques recourent à la transcription surtout pour les manuscrits qui posent des problèmes au niveau de leur océrisation rendant difficile leur mise à disposition en mode texte. Des transcriptions collaboratives peuvent jouer alors un rôle important dans la résolution de ces difficultés. 35% des corpus d'archives numériques utilisent le module « annotation collaborative » pour donner l'occasion aux internautes d'indexer dans un état civil par exemple, une année, un nom et un prénom dans le but de créer des points chronologiques (années) et un réservoir de noms pour aider à la recherche historique.

En contrepartie, 50% des services d'archives numériques adoptent la pratique de folksonomie en comptant sur l'investissement et le volontariat des usagers dans les travaux d'identification des archives. En effet, ils mettent à la disposition des internautes la possibilité d'indexer des documents et surtout des photographies à l'aide des tags (mots clés) choisis librement. Cette identification facilite l'accès à ces photographies par les moteurs de recherche

Contrairement à la folksonomie, 35% des corpus d'archives numériques recourent à l'indexation collaborative qui fait appel à un encadrement plus strict (Néroulidis, 2015) en obligeant l'internaute à l'inscription dans un compte personnel avant de participer à l'indexation, et en faisant un test de paléographie au moment du travail sur des lots choisis par les archives (Les Archives départementales de l'Aube, les Archives départementales Puy de Dôme).

IV- Discussion :

Les humanités numériques avec leurs outils et pratiques innovants constituent une véritable révolution dans l'histoire de la mémoire de l'humanité en ce qu'elles facilitent et améliorent sa transmission. Avec les technologies numériques qu'elles ont introduites, les archives ont subi des transformations majeures touchant à la fois leur notion qui s'ouvrent aux « corpus », « collections numériques », leurs pratiques de traitement et de conservation à long terme et leurs nouvelles fonctionnalités collaboratives qui s'ouvrent à la notion des archives participatives.

En contrepartie des remarques qui touchent les principes archivistiques, les instruments de recherche mis en ligne, les nouvelles fonctionnalités collaboratives induisent ces questionnements méthodologiques :

1. Les principes archivistiques à l'ère des humanités numériques est-ce un « paradoxe » ?

Selon Yann Potin (Potin, 2011, p.5), la mutation vers le numérique a engendré la confrontation des archives avec les nouveaux modes d'accès aux archives numériques à l'égard de l'utilisation des moteurs de recherche plein texte, ou les thesaurus aux mots clés tels que les corpus d'archives de Val de Marne qui utilisent le thésaurus W et le thésaurus MOTBIS ainsi que d'autres corpus.

Ces modes d'accès aux contenus de corpus d'archives se rapprochent ou reproduisent le modèle du catalogue de bibliothèque (Potin, 2011, p.5) modèle basé sur la réduction thématique supposée par la logique d'indexation, ce qui provoque une nouvelle confrontation entre deux modes de classement (le principe de pertinence et le principe de provenance)

2. Vers la fin des instruments de recherche ?

Le numérique avec ses méthodes actuelles de mise en ligne, amorce une rupture avec la culture de l'imprimé (politiques éditoriales, identification matérielle et intellectuelle) et avec toute la culture scientifique moderne (Coulon, 2010, p.10). Cette nouvelle opportunité de diffusion des documents (la mise en ligne) a tendance à se réaliser sans rapport avec la culture de l'imprimé précédemment exposée, et c'est là, pour nous, que réside une partie du problème (Coulon, 2010, p. 26). Selon l'auteur, cette ignorance de la culture de l'imprimé pose des problèmes au niveau des procédés de publication et de présentation des instruments de recherche en ligne quand leur structuration et leur affichage de l'information n'est pas accessible au lecteur (situation nécessitant le besoin de déplier toute la hiérarchie de l'instrument et effectuer des requêtes).

Louis Colombanie en s'inspirant de Baptiste de Coulon parle de la fin ou de la quasi-disparition des instruments de recherche du fait de leur mise en ligne ou leur publication sous forme numérique. Il montre qu'à l'ère numérique, les instruments de recherche mis en ligne par quelques services d'archives accessibles en ligne risquent de perdre leur statut de documents de nature scientifique publiés (dans certaines archives départementales en ligne, les instruments de recherche sont bien présents, mais sont sous forme d'un fichier PDF difficilement lisible). En effet, les instruments de recherche actuels mis en ligne ne traduisent pas leur rôle effectif de guider le chercheur ou de faciliter l'accès au document recherché. Selon Louis Colombanie, ce qui existe ne sont pas des instruments de recherche mais des « arbres » qui se déplient laborieusement, qui vous repositionnent malgré vous tout en haut quand vous pensez avoir enfin réussi à atteindre la description qui vous intéresse.

Conclusion

Les archives sont la mémoire de l'humanisme. Elles résument ses forces et ses faiblesses, ses souvenirs et ses oublis, ses arts et ses littératures. Avec les avancements numériques, les Archives ont joué encore plus ce rôle, justifiant à l'ère des humanités numériques une mémoire dite numérique.

Suivant notre étude d'évaluation effectuée sur 20 corpus d'archives numériques, nous avons constaté que les humanités numériques font évoluer les archives au niveau de leurs outils, leurs fonctionnalités archivistiques et leurs méthodes de conservation à long terme.

L'interface en tant qu'outil de présentation de contenu joue un rôle important dans les projets orientés vers les humanités numériques, ce qui justifie que la totalité généralement (100%) des corpus d'archives numériques utilisent des interfaces plus ergonomiques, modernes et attractives dans la mesure où leur présentation visuelle est attractive et leur navigation est satisfaisante.

Outre l'interface ergonomique, les corpus d'archives numériques évalués présentent un contenu évolué au niveau de sa crédibilité, son interactivité et ses fonctionnalités archivistiques.

En effet, 87% des corpus d'archives numériques sont qualifiés fiables au niveau de l'information proposée (présence de références, copyright, témoignages...). 100% des corpus d'archives respectent la responsabilité intellectuelle (mention des noms, logos, références de responsables...). Entre 90% et 95% des corpus mentionnent la source de leur financement et pour le référencement presque la totalité des corpus d'archives sont référencés de la même façon par les moteurs de recherche en suivant des recommandations internationales d'accessibilité et en présentant un accès libre aux archives. Cet accès autorise la réutilisation des informations contenues dans leurs corpus à condition de ne pas en faire un usage commercial des données, de ne pas les diffuser publiquement ou à des tiers à l'exception d'un usage privé ou interne.

Les fonctionnalités archivistiques elles-mêmes ont connu une évolution remarquable :

- Au niveau de la description par les métadonnées qui intervient dès la création du document numérique par le moyen des métadonnées et des structures d'encodage (EAD, EAC,) : 76% des corpus utilisent l'EAD
- Au niveau de la recherche et de la présentation des résultats : 100% des corpus utilisent la recherche par mots clés, et la majorité utilise la recherche par collection spécialisée.

- Au niveau des outils de lecture développés édifiés par l'utilisation d'une interface développée de consultation de documents : 80% utilisent l'outil de visionnement des documents *numérisa*).
- Au niveau du travail collaboratif : entre 35% et 40% participent aux activités de transcription, d'annotation et d'indexation collaboratives. 65% des corpus sont signalés sur des réseaux sociaux et professionnels.

Le numérique a changé les pratiques de conservation des archives en apportant de nouvelles perspectives dans la notion de conservation à long terme des archives qu'il s'agissent des formats utilisés, ou des techniques de permaliens ou des techniques de moissonnage (44% des corpus utilisent le PDF, et 24% utilisent la technique de permaliens).

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 plus adaptées à créer de nouvelles tendances, et les archivistes doivent dépasser le rôle de gardiens de la mémoire pour accéder à celui de contrôle et de gestion.

Peut-on parler encore d'un rôle de créativité de l'archiviste passant des archives aux données massives créées par des *big data*, et des humanités numériques vers encore plus d'humanités créatives dans le cadre des projets de type *crowdsourcing* ?

Un autre volet pourrait être développé suite à cette recherche, celui de parler des humanités numériques dans le monde arabe qui est caractérisé par son histoire culturelle et linguistique plus riche et par son patrimoine traversé par des diversités civilisationnelles (berbère, punique, romaine, carthaginoise, chrétienne, arabe) et coloniales (française, espagnole, britannique). La question est de savoir comment et quand le monde arabe investira dans les enjeux des humanités numériques pour garantir l'accès et le partage des ressources du savoir ?

Références bibliographiques

1- Banat-Berger, Françoise, Laurent Duplouy et Claude Huc. 2009. L'archivage numérique long terme les débuts de la maturité, Paris: la Documentation française.

2- 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.).

3- Burnard, Lou.,2012. « Du literary and linguisticcomputing 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. OpenEditionPress, 2012. http://books.openedition.org/oep/242.consultable en ligne :

4- 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

5- 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>

6- 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 PhilosophiæDoctor (Ph.D.) en sciences de l'information : Université de Montréal

7- Clément ,Jean, 2001. La littérature au risque du numérique, http://www.cairn.info/article.php?ID_ARTICLE=DN_051_0113

8- Couture, Carol, Jacques ducharme et Jean-Yves Rousseau, 1988. L'archivistique a-t-elle trouvé son identité? Argus, vol 17,n2,p.51-58

9- 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

10- Danaks Patricia, 2014. Question interdisciplinaire : la philosophie et l'archives, https://www.academia.edu/10655438/Question_interdisciplinaire_la_philosophie_et_larchive._Automne_2014

11- 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

12- Delpierre, Nicolas ; Delpierre,Françoise ; Mirguet, Françoise. 2012. Les chantiers du numériques : Dématérialisation des archives et métiers de l'archiviste. Paris : Academia.

13- 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

14- Duranti, Luciana. 1995. Archives as a Place. In Paper Presented at a Half-Day Seminar in Sydney.

15- Gagnon-Arguin, Louise ;Vien, Hélène. 2002. l'archivistique contemporaine : Formation continue à l'usage des fonctionnaires tunisiens. Tunis : Archives Nationales de Tunisie.

16- Kenneth ,M.Price, 2009 "Edition, project, Database, Archive, Thematic Research Collection :what's in a Name? Digital Humanities Quarterly 3.3. http://www.digitalhumanities.org/dhq/vol/3/3/000053/000053.html
17- Le Deuff, Olivier.2012 Humanisme numérique et littératies. Semen - Revue de sémio-
18- linguistique des textes et discours, Presses Universitaires de l'Université de Franche Comité (Pufc), 2012, pp.117-134. http://archivesic.ccsd.cnrs.fr/sic_00776805v1/document
19- Le Deuff, Olivier. 2015. Les humanités digitales précèdent-elles le numérique ? . Imad Saleh. H2PTM 15, Iste éditions, 978-1784051150.
20- 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/#/
21- 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
22- Mourier ,Jacques; Caya, Marcel , 2011.module 1:les archives pour quoi?
23- Müller, Bertrand 2011 .Archives et temps présent : considérations inactuelles ; Communication présentée au colloque Temps présent et contemporanéité, Paris, 24-26 mars 2011
24- Pastinelli, Madeleine , 2009 "La mémoire et l'oubli dans l'univers de l'archive totale." Espaces Temps.net, Travaux, 19.02.2009. http://www.espacestemps.net/articles/archive-totale/
25- Poivre, Joel, 2004. L'archivage électronique : principes, promesses et pièges, http://adeda78.free.fr/IMG/pdf/POIVRE_promesses.pdf
26- Portier, Pierre-Édouard, 2010. Construction des documents multi structurés dans le contexte des Humanités numériques, file:///C:/Users/fatma/Desktop/les%20humanités%20numériques/a%20lire/Liris-4893.pdf
27- Schellenberb, Theodore.1956. Modern archives: Principles and Techniques; University of Chicago
28- 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
29- 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
30- Vitali Rosati, Marcello, 2011. Une philosophie du numérique lecture pour un humanisme numérique de Milad Doueihi. Paris : Seuil.
31- Viiri, Sampo. (2014). Digitl Humanities and Future Archives (reports). London. Finnish Institute in London.2014. 52p. [en ligne] http://www.fininst.uk/wp-content/uploads/2017/09/Digital_Humanities_and_Future_Archives.pdf