



Diplôme national de master

Domaine - sciences humaines et sociales

Mention - sciences de l'information et des bibliothèques

Spécialité - archives numériques

L'intelligence collective au service de la valorisation des archives numériques publiques

Sokhna Gueye

Sous la direction de Céline Guyon
Maître de conférence - ENSSIB



Remerciements

J'adresse mes remerciements les plus sincères à toutes les personnes qui ont contribué de près ou de loin à l'élaboration de ce mémoire.

Je tiens à exprimer en premier lieu toute ma reconnaissance à Céline Guyon, ma directrice de mémoire pour sa disponibilité et son écoute incommensurables.

Je remercie profondément Jean-Yves Leclerc qui m'a beaucoup aidé dans l'élaboration de mon enquête.

Mes pensées vont également à l'endroit de ma famille, pour leur écoute et leur soutien, particulièrement mon époux qui m'a beaucoup soutenu durant la rédaction de ce mémoire.

Enfin je tiens à remercier mon amie Aida Tamboura pour son temps qu'elle m'a accordé pour la relecture et les corrections.

Résumé :

L'évolution du Web implique l'apparition de nouvelles pratiques participatives reposant de l'intelligence collective. Les opérations de crowdsourcing apportent beaucoup à la valorisation des archives dans les services publics. Les internautes, sollicités pour enrichir les fonds, apportent aussi de la valeur ajoutée sur les archives.

Descripteurs :

Intelligence collective, crowdsourcing, web social, archives numériques, valorisation

Abstract:

The evolution of the Web implies the emergence of new participatory practices based on collective intelligence. Crowdsourcing operations bring a lot to the value of archives in public services. Internet users, asked to enrich the funds, also bring added value to the archives.

Keywords :

Collective intelligence, crowdsourcing, social web, digital archives, valorization.

Droits d'auteurs

Cette création est mise à disposition selon le Contrat : « Paternité - Pas d'Utilisation Commerciale - Pas de Modification 2.0 France » disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/3.0/deed.fr> ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Sommaire

Sommaire	5
Introduction	7
I. ÉLÉMENTS DE TERMINOLOGIE	9
A. L'intelligence collective	9
1. L'intelligence collective à travers l'analogie de la fourmilière	9
2. La sagesse des foules	12
3. La puissance d'agir des sujet sociaux	14
B. Le contributeur et la contribution en ligne.....	15
C. Le <i>crowdsourcing</i>	18
1. Définitions diverses, sens convergents	18
2. Participation involontaire des internautes.....	19
3. Origines historiques et conceptuelles du <i>crowdsourcing</i>	20
4. Typologie du <i>crowdsourcing</i>	25
D. Autres notions liées au <i>crowdsourcing</i>	26
1. Le Web 2.0	26
2. L'indexation sociale (métadonnées sociales) ou <i>folksonomies</i> ou indexation communautaire	28
3. Les « archives participatives »	29
II. CROWDSOURCING DANS LES SERVICES D'ARCHIVES	32
A. Enquête auprès des services d'archives ayant mis en place des projets de <i>crowdsourcing</i> : méthodologie adoptée	32
1. Conception des outils d'enquête	33
2. Présentation de la population et de l'échantillon de l'enquête.....	33
3. Les services d'archives et institutions répondants.....	36
B. Panorama du <i>crowdsourcing</i> dans les services d'archives.....	39
1. Ajout de documents, de photographies, de cartes postales.....	40
2. Identification de documents iconographiques	41
3. L'annotation collaborative	44
4. Indexation collaborative	46
5. Transcription collaborative	51
6. Edition collaborative de carnets de guerre.....	53
7. Catégorisation des projets	54
III. LA VALORISATION DES ARCHIVES ET LE CROWDSOURCING	62
A. Qu'est-ce que la valorisation des archives ?.....	63

1. La valorisation dans la chaîne archivistique	65
2. Diffuser et valoriser : quelles différences ?	67
3. Forces des services d'archives : masse documentaire, documents uniques et aux valeurs multiples, etc.	69
B. La valorisation des archives par le <i>crowdsourcing</i> : comment ?.....	71
1. Évolution des pratiques du <i>crowdsourcing</i> : quels impacts sur la mise en valeur des archives ?.....	73
2. Quels projets pour valoriser quels fonds ?	75
3. Les publics des projets : types et motivations	80
4. Les moteurs de réussite des projets et de la valorisation des fonds	82
5. Les finalités des projets de <i>crowdsourcing</i> sur les fonds	91
Conclusion	94
Sources	95
Bibliographie	96
Table des annexes.....	100
Entretien	113
Table des illustrations.....	114
Table des matières.....	116

NOTA BENE : Dans ce travail, nous avons utilisé des citations d'auteurs qui sont parfois mis entre crochets. Les citations ne sont pas identiques aux originales ; elles sont commentées et reformulées. Mais nous avons pris le soin de citer les auteurs à chaque fin de phrase ou en note de bas de page.

Introduction

Inventé par Tim O'Reilly¹, le web 2.0 ou web social ou collaboratif met l'accent sur une nouvelle forme d'interactivité qui place l'utilisateur au centre de l'Internet. L'implication des utilisateurs dans le réseau constitue le facteur clé dans cette vision du web. Son importance est mise en valeur par le phénomène des blogs, des wikis et le développement des *folksonomies*.

Avec le web 2.0 de nouvelles pratiques collaboratives comme le *crowdsourcing* font leur apparition. Aujourd'hui il est de plus en plus pratiqué dans les domaines culturels et patrimoniaux. Plus particulièrement dans les services publics ces nouvelles pratiques du Web 2.0 permettent d'initier de nombreuses activités avec les usagers autour des fonds.

Les archivistes de leur côté doivent mettre en valeur leurs fonds. Ainsi, ils sont à la recherche d'outils pour inventer de nouvelles formes de médiation des fonds, pour paraphraser Elodie Belkorchia². Ces nouvelles formes de valorisation passent désormais par le Web qui offre de nombreuses possibilités aux archivistes. Ces facilités se concrétisent par des pratiques organisées sous forme de projets qualifiés de participatifs ou encore de collaboratifs en fonction du degré d'implication de l'internaute. Ces projets vont de l'indexation collaborative à la transcription de textes en passant par l'annotation et l'édition de documents (carnets de soldats). Par l'intermédiaire des internautes, ces projets apportent de la valeur ajoutée aux fonds d'archives. Cette plus-value est justifiée par la qualité des données produites par les internautes qui sont souvent réutilisées par les services d'archives.

Existe-t-il donc un lien entre la valorisation des fonds d'archives et ces projets participatifs qui se multiplient dans les services d'archives ? D'où notre sujet dont la question centrale est « comment l'intelligence collective participe-t-elle à la valorisation des archives numériques publiques ? ». Le *crowdsourcing*, ainsi que les pratiques dérivées de l'intelligence collective ont fait l'objet de plusieurs études dans l'univers des archives. Par ailleurs l'aspect valorisation des fonds à travers ces pratiques participatives est très peu soulevé dans la littérature scientifique. Ce sujet vise donc à étudier les formes d'intelligence collective, plus particulièrement le *crowdsourcing* sous l'angle de la valorisation des archives numériques publiques. Dans ce cadre nous précisons que nous nous intéressons aux services publics d'archives français (communes, départements, régions, ministères, et autres organismes

¹ O'REILLY, Tim. What is Web 2.0. *O'Reilly Media* [en ligne]. 2005. Disponible sur : <<http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>> (Consulté le 7/12/2017).

² La Gazette des Archives, *Opt. Cit.*

publics). Tout au long de ce mémoire nous parlerons d'« archives numériques » pour les « archives numérisées » sachant qu'elles ne sont pas nativement numériques. De même le *crowdsourcing* sera mis en avant pour signifier une forme d'intelligence collective et considéré comme une pratique contributive et collaborative. Ainsi, nous utiliserons la notion de « projets collaboratifs » et de « projets participatifs » pour parler des « projets de *crowdsourcing* ». Car nous considérons que ceux-ci sont aussi bien collaboratifs que participatifs, du fait de l'implication de certains contributeurs.

Afin d'apporter des éléments de réponse à notre questionnement, nous allons dans un premier temps entamer une réflexion autour de l'intelligence collective, faire un panorama de définitions apportées par les spécialistes du sujet. Nous ferons un focus sur le *crowdsourcing*, en essayant de remonter à ses origines historiques et conceptuelles et en introduisant les auteurs qui ont récemment étudié ce sujet. Toujours dans la première partie, nous allons apporter des définitions de quelques concepts, trouvés dans la littérature scientifique, et qui souvent, se rapportent à l'intelligence collective et spécifiquement au *crowdsourcing*.

Après cette phase de définition nous ferons un état des lieux sur les projets participatifs dans l'univers des archives. L'objectif étant de voir la diversité des projets et des pratiques afin de mieux cerner la question. Nous nous appuyons essentiellement sur ces projets pour constituer la base de l'échantillon de notre enquête auprès des services d'archives. Ainsi, nous présenterons notre méthodologie d'enquête, la population et l'échantillon, ainsi que les services d'archives répondants. Pour cela l'élaboration des matériaux sera exposée ainsi que les premiers résultats obtenus.

Enfin nous essaierons de démontrer comment l'intelligence collective participe-t-elle à la valorisation des archives numériques publiques. Ainsi, nous étudierons ce qu'est la « valorisation des archives » afin d'exposer son importance dans la chaîne archivistique. Pour répondre à notre problématique, nous aborderons l'évolution des pratiques du *crowdsourcing* pour ainsi percevoir les liens avec les stratégies de valorisation. Nous nous intéresserons aussi aux différents types de publics, leurs rôles et leurs apports dans cette valorisation. Les finalités des projets collaboratifs seront abordées afin de voir leurs retombées sur les fonds. Mais avant cela nous parlerons des moteurs de réussite des projets qui permettent de valoriser mieux les fonds. La question de la réutilisation de la production des usagers dans le cadre de ces projets sera aussi étudiée ainsi que son rôle dans la mise en valeur.

I. ÉLÉMENTS DE TERMINOLOGIE

L'intelligence collective et les concepts liés à ce phénomène ont fait couler beaucoup d'encre dans la littérature scientifique. Phénomène, comme le nomment beaucoup d'auteurs, l'intelligence collective est un sujet très complexe. Dans cette partie, nous apporterons des éléments de définition à ce phénomène, ainsi que les concepts lui sont liés.

A. L'intelligence collective

L'intelligence collective ou l'intelligence du grand nombre ou encore l'intelligence des foules, autant de concepts attribués à ce phénomène dont il est difficile d'identifier ses origines. Entamer une réflexion sur l'intelligence collective paraît une tâche difficile puisque les spécialistes de la question n'arrivent pas à mettre d'accord sur son existence. En effet ce phénomène est un sujet à détractations. Bon nombre d'auteurs contredisent ceux qui ont voulu octroyer une intelligence aux foules. Pour Serge Proulx, « prétendre qu'il existe une « intelligence du grand nombre » apparaît à première vue paradoxal³. La foule est du côté du grand nombre et vouloir attribuer une intelligence au grand nombre apparaît tâche difficile. » Ainsi, il s'interroge sur le fait d'instaurer le grand nombre comme le siège d'une intelligence collective alors que la foule apparaît comme disqualifiée en matière de formulation ou de défense efficace d'opinions argumentées susceptibles de fonder des actions publiques pertinentes [Proulx, 2009, p. 9].

Les sociologues et autres spécialistes de la question ont usé diverses manières pour expliquer ce phénomène de l'intelligence des foules. L'une de ces tentatives de définitions qui a attiré notre attention est l'analogie entre l'intelligence collective et les sociétés d'insectes.

1. L'intelligence collective à travers l'analogie de la fourmilière

« L'intelligence collective serait une « fourmilière (ou tout autre société d'insectes) » où chacun réagit à la façon de l'autre ... »

³ HASSOUN, Mohamed, EL HACHANI, Mabrouka. *Intelligence collective et organisation des connaissances : 7^e colloque international du chapitre Français de l'ISKO, Lyon, 24-26 juin 2009*. Lyon : Université Jean Moulin Lyon 3, 2009, 399 p.

Cette analogie démontrée par Ferber, J.⁴, permet de comprendre le fonctionnement de la pratique de l'intelligence collective. Les êtres humains participants apportent leur part de l'intelligence constituée par le groupe. Cette intelligence collective permet le plus souvent d'apporter des solutions à un problème constaté. L'analogie entre l'intelligence collective et les sociétés d'insectes suppose que les individus agissent comme les insectes à la recherche de nourriture. Dans le même ordre d'idée, Joël de Rosnay affirme ce fait dans la préface de *La sagesse des foules*⁵, en avançant :

« Les réseaux d'individus connectés à Internet sont capables de réactions rapides et d'adaptations. On observe un phénomène d'essaims (*swarm intelligence*) comme ceux formés par les abeilles et les termites. ... Sur Internet on pourrait appeler de tels regroupements humains, temporaires ou permanents, des "cyber essaims " ». [p. 11]

Par le biais de cette analogie, James Surowiecki, tente de démontrer plutôt qu'il existe bien une intelligence du grand nombre dans ses expériences étudiées dans l'ouvrage précité. Par l'exemple des colonies d'abeilles, il montre comment celles-ci parviennent à trouver une solution collective au problème de nourriture. De son point de vue les abeilles sont remarquablement efficaces pour trouver de la nourriture en cherchant la source à six kilomètres, et s'il y a un parterre de fleurs à deux kilomètres de l'essaim. Leur méthode est la suivante, elles envoient un groupe en éclaireur pour prospecter les alentours. Quand un éclaireur trouve une source de nectar qui semble de bonne teneur, il revient et se livre à une danse dont l'intensité dépend, en quelque sorte, de l'excellence de la source de nectar du site. La danse attire les autres éclaireurs abeilles, qui suivent la première. Par ailleurs les éclaireurs qui ont trouvé des sites moins bons attirent moins d'abeilles dans leur sillage et, dans certains cas, finissent par abandonner leur site. Au final, les éclaireurs abeilles se répartissent de manière quasi parfaite sur les différentes sources de nectar, ce qui veut dire qu'elles accumulent autant de nourriture possible en fonction du temps et de l'énergie accordés à la recherche. La chose importante à retenir dans cet exemple est la façon dont la colonie parvient à une solution collective intelligente. Ainsi, l'auteur confirme :

« Elle n'y parvient pas en examinant de manière rationnelle toutes les alternatives et en élaborant un mode idéal de fouissage. Elle ne peut pas le faire parce qu'elle n'a aucune idée des alternatives possibles, c'est-à-dire des différents parterres de fleurs. Elle envoie donc des

⁴ FERBER, J. *Les systèmes multi-agents : vers une intelligence collective, informatique, intelligence artificielle*. Paris : InterEditions, 1997.

⁵ SUROWIECKI, James ; RIOT, Elen (trad.). *La sagesse des foules = The Wisdom of crowds*. Paris : Editions Jean-Claude Lattès, 2008, 384 p. ISBN : 978-2-7096-2891-4

éclaireurs dans toutes les directions avec l'espoir qu'au moins l'un d'entre eux finira par trouver le meilleur parterre, revenir, et faire une bonne danse qui informera l'essaim du lieu de la source de nourriture ». [p. 62]

Cet exemple montre bien que l'intelligence d'un groupe a permis de solutionner collectivement un problème. Les colonies d'abeilles sont parvenues à trouver leur nourriture sans tenir de réunion et utiliser une technique très efficace.

Par ailleurs l'analogie de la fourmilière, conçoit l'être humain comme un participant neutre qui n'a aucun pouvoir ni contrôle sur son action encore moins sur les connaissances produites. C'est dire qu'à l'image des sociétés d'insectes, chaque être humain imite l'autre et agit exactement comme lui. Alors ne pourrait-on pas se poser la question : existe-il vraiment une intelligence du grand nombre ? Car la plupart des critiques récurrentes vis-à-vis de l'intelligence collective ont tendance à nier son existence en affirmant que « le grand nombre n'est pas clairvoyant ; il est plutôt stupide, incapable de produire des connaissances fondées, valables ». Ainsi, Mayere⁶ accentue ces réprobations à l'égard de l'intelligence collective en avançant que cette dernière n'a d'intelligent que de nom et ne servirait qu'à centraliser la mémoire pour mieux la contrôler sans permettre l'émergence de nouvelles connaissances. Pendant ce temps d'autres sociologues vont jusqu'à démontrer l'inexistence d'une intelligence des foules.

Dans l'ouvrage collectif *Intelligence collective et organisation des connaissances* publié à l'issue du 7^e colloque international du chapitre français de l'ISKO, Serge Proulx⁷ met en exergue les propos de Gustave Le Bon. Ce dernier, dans sa *Psychologie des foules* en 1895 affirmait que les foules agissaient comme des barbares incapables de réflexion, subissant l'influence des images plutôt que des arguments rationnels, mues par un inconscient collectif incontrôlable, susceptibles de comportements contrastés et variables, parfois même violents.

Cette opinion peut être discutable dans la mesure où l'intelligence collective permet non seulement de produire des connaissances nouvelles mais aussi avec ou sans l'analyse et la combinaison de données, il peut y avoir une émergence d'autres connaissances. Selon Szoniacky cette conception de l'intelligence collective a tendance à faire de l'humain participant un être

⁶ MAYERE, A. « L'intelligence collective : une notion en chantier », *Intelligence collective*. : Paris : Presses de MINES, 2006

⁷ Proulx in *Intelligence collective et organisation des connaissances, Opt. Cit*, p. 9

simple et basique, dont il faut soustraire toute complexité pour faire réagir chacun à la manière de tout le monde [Szoniecky, 2017, p.13]. Dans cette optique, il avance :

« L'intelligence collective est donc la mise en commun de la réflexivité afin de finaliser une action qui ne pourrait pas l'être par une seule personne. » [p. 14]

Pour lui, l'intelligence collective est loin d'être analogue à une société d'insectes car derrière la centralisation de la connaissance, il existe bien une réflexivité qui différencie d'ailleurs l'homme de la fourmi. Car il a la capacité de réfléchir à la mise en commun contrairement à la fourmi. Parallèlement pour LEVY⁸, ce qui différencie l'être humain de la fourmi c'est sa capacité à réfléchir les flux d'informations dans son intériorité et ainsi exprimer une conscience réflexive. Ainsi, Yves Prié précise que la réflexivité est la faculté de rentrer en relation avec soi-même pour construire à partir de sa mémoire des représentations permettant la régulation de son action⁹.

2. La sagesse des foules

L'émergence d'Internet et, en particulier l'usage des plateformes collaboratives, a pourtant remis à l'ordre du jour l'hypothèse d'une « intelligence des foules » selon James Surowiecki. En 2004, il publie la *Sagesse des foules (The Wisdom of Crowds)* pour démontrer à travers des expériences l'intelligence des foules. Dans ce cadre il définit la foule comme étant « tout groupe susceptible d'agir collectivement dans la prise de décision ou la solution de problèmes » (grandes organisations privées ou gouvernementales, mais aussi regroupements plus restreints, comme par exemple un réseau de scientifiques). Surowiecki inclut également dans son corpus des agrégats d'individus en action qui ne sont pas conscients qu'ils forment une entité collective, par exemple : des parieurs sur un champ de course ou des investisseurs dans le marché financier. Dans la préface de cet ouvrage, Joël de Rosnay revient sur les débuts de l'intelligence des foules avec Eli Lilly. Cette dernière est une grande entreprise pharmaceutique qui avait lancé une start-up et s'était appuyée sur un réseau interconnecté de 80 000 chercheurs indépendants appartenant à 170 pays pour résoudre des problèmes posés à la Recherche du groupe et à certains de ses partenaires. Une intelligence des foules était en train de naître sans qu'on puisse la nommer ainsi jusqu'en 1995 dans *L'homme symbiotique*. Avec cette expérience l'auteur démontre l'existence d'une intelligence des foules qui, connectées en temps réel, peuvent résoudre des problèmes complexes, parfois mieux que le meilleur des experts. Dans

⁸ Intelligence collective et archives numériques, *Opt. Cit.*, p. 13

⁹ Intelligence collective et archives numériques, *Opt. Cit.*, p. 13

son ouvrage, Surowiecki tente de démontrer cette sagesse collective à travers différentes expériences. Le travail étant d'examiner une série de façons différentes de puiser dans le savoir d'un groupe : prix des actions, votes, la cote des paris mutuels, algorithmes informatiques et contrats à terme. Ce sont autant de tentatives pour puiser dans la sagesse des foules, et c'est la raison pour laquelle elles fonctionnent.

À travers ces expérimentations, il a pu montrer que l'intelligence du groupe a permis de trouver des solutions à des problèmes parfois avec ou sans concertation. Dans tous les cas, l'intelligence du groupe a été utilisée pour découvrir une ou des alternatives possibles pour solutionner un problème. La foule arrive à trouver des solutions certes mais il faut aussi se poser la question à savoir est-ce les bonnes décisions ? Parmi les exemples présentés par l'auteur il semble que les groupes sont aptes à distinguer les bonnes solutions des mauvaises. Cela est possible grâce à la diversité des membres du groupe.

Surowiecki se questionne aussi sur l'hétérogénéité d'un groupe. Ainsi, peut-on dire que la diversité influe sur un groupe ? Autrement dit, une fois qu'un éventail de solutions possibles a été élaboré, est-ce que le fait d'avoir un groupe de décideurs varié fait la différence ? Il insiste sur le fait que la diversité peut aider le groupe dans la mesure où elle augmente ses chances de trouver la bonne décision et diminue les aspects destructeurs qui empêchent d'aller vers la bonne direction. Ainsi, il soutient :

« La diversité est une aide au sens où elle accroît les perspectives, et parce qu'elle annule, ou du moins affaiblit, certaines des caractéristiques destructives de la prise de décision du groupe. L'importance de la diversité cognitive ne signifie pas qu'un simple groupe de personnes diverses mais dépourvues de la moindre information offrira une sagesse collective supérieure à celle d'un expert. Mais assemblez un groupe hétérogène de personnes douées de savoir et de perspicacité à des degrés divers ; et vous pouvez davantage lui faire confiance pour les grandes décisions qu'à une ou deux personnes, aussi qu'intelligentes qu'elles soient. Si cela semble difficile à croire [...] c'est parce que cela va à l'encontre de nos institutions fondamentales à propos de l'intelligence et des affaires ».

La sagesse des foules se manifeste à travers leur puissance d'agir. En effet les sujets sociaux sont capables d'agir en nombre notamment sur les réseaux sociaux numériques. Les foules ont non seulement une intelligence mais elles ont aussi une puissance d'agir qui leur permettent d'avoir une influence par exemple sur le vote pour certains candidats lors d'une élection selon Surowiecki.

3. La puissance d’agir des sujet sociaux

La notion d’intelligence du grand nombre est souvent liée à la puissance d’agir des sujets sociaux qui elle, est rattachée aux concepts d’*empowerment* et d’*agency* d’après Serge Proulx et Jérôme Vidal¹⁰. Pour ce dernier, l’ « *agency* » se traduit comme la capacité d’agir des acteurs sociaux. Cette *capacité d’agir* aurait l’intérêt de dresser une passerelle entre la problématique de l’*agency* et les sociologies de la domination qui, comme celle de Pierre Bourdieu, posent que la capacité statuaire (légitime, reconnue, officielle) conditionne le développement de la capacité (affective, psychique, intellectuelle, physique ...) effective - mais ce serait rabattre la problématique de l’*empowerment* sur celle, institutionnelle, de l’habilitation » [Vidal, 2010].

La puissance d’agir des sujets sociaux est favorisée aujourd’hui par les outils du Web 2.0. En effet le Web constitue un moyen facile qui permet d’agir en grand nombre. Certains auteurs ont pu constater que les manifestations sur le web s’accroissent parfois lors d’évènements majeurs. C’est le cas des attentats de Paris notamment au Bataclan en novembre 2015 qui ont fait l’objet de beaucoup de réactions sur Twitter et Facebook et sur d’autres réseaux sociaux. Chez Proulx, les sujets sociaux sont souvent inconscients de cette puissance d’agir qu’ils détiennent. Dans ce cadre il affirme :

« Chacun est entouré de personnes qui ont les mains plongés dans l’univers Internet, qui « manipulent » ... sans le savoir bien souvent, une certaine « puissance d’agir » à travers le développement d’une « culture de la contribution », qui représente une force de résistance face à l’emprise d’un capitalisme informationnel et cognitif. »¹¹

En effet la culture de la contribution s’est installée avec les débuts du Web dans les années 2000 et est surtout suscitée par de nombreux facteurs comme le développement du logiciel libre, les *Requests for comments (RFC)*, suggestions de modifications, des bricolages informatiques, etc. [Proulx, 2014]

Avec la montée en puissance des technologies de l’information et de la communication, nous pouvons constater que la force d’agir des sujets sociaux est mise en exergue par la vie en ligne des internautes. Cette vie ligne est partagée entre le temps passé sur les réseaux sociaux sur les comptes personnels, la recherche sur le web, le shopping et les pratiques participatives. Cela nous amène à nous intéresser d’abord sur la contribution en ligne avant d’amorcer la réflexion

¹⁰ VIDAL, Jérôme. *Jérôme Vidal, le blog de... Jérôme Vidal : articles, entretiens, traductions et autres trucs*. [En ligne]. (Publié en 2010) Disponible sur : <<http://jeromevidal.blogspot.com/2010/07/la-question-de-lagency-puissance-et.html>> (Consulté le 08/07/2018)

¹¹ GARCIA, José Luis, HEATON, Lorna, PROULX, Serge, dir. *La contribution en ligne : pratiques participatives à l’ère du capitalisme informationnel*. Québec : Presses de l’Université du Québec, 2014, 256 p.

sur le *crowdsourcing*. Cela nous permettra de comprendre la logique de la gratuité dans cette culture de la contribution et le positionnement du contributeur.

B. Le contributeur et la contribution en ligne

Avec l'automatisation et la transformation numérique, les conditions du travail humain s'en voient aujourd'hui totalement ou partiellement transformées. Dès lors « une économie de la contribution » s'installe et conquiert beaucoup de domaines. Ainsi, nous sommes témoins d'un nouveau modèle industriel. Selon Serge Proulx ce phénomène c'est un troisième modèle industriel où la figure hybride de l'utilisateur contributeur s'impose. Ce dernier n'est ni exclusivement un producteur ni seulement un consommateur [Proulx, 2014, p. 16]. Son statut se situe entre le producteur et le consommateur. Il incarne alors la figure du « producteur-utilisateur » (*producer*) selon l'expression d'Axel Bruns [Proulx, 2014]. Alors le contributeur est :

« Un amateur au vieux sens du terme. C'est d'abord quelqu'un qui est motivé par ses centres d'intérêt plutôt que par des raisons économiques. Il peut d'ailleurs développer une expertise plus grande que ceux qui sont motivés par des raisons économiques » [Proulx, 2014, p. 16].

Dans ce nouveau modèle du travail, le contributeur apporte davantage que sa force de travail : il apporte du savoir, des connaissances. Cette analyse orientée vers le rôle de la connaissance dans la production de la valeur économique et les développements contemporains de l'industrie, est en phase avec les travaux d'économistes réalisés depuis 2004 autour de l'hypothèse de l'émergence d'un « capitalisme cognitif » [Proulx, 2014, p. 17].

Ce modèle du travail se répercute dans le monde de l'information. En effet la relation entre le professionnel et l'utilisateur change complètement.

« Ce dernier n'est plus seulement un consommateur, mais un acteur ; ou, pour mieux dire, en utilisant un mot-valise français cette fois, un «consommateur» de la connaissance, qui enrichit volontairement l'instrument de recherche ou le document d'archives numérisé lui-même » [Édouard Bouyé, 2012, p. 126]¹².

La contribution en ligne s'est beaucoup développée avec l'émergence de l'Internet et des réseaux sociaux numériques. En effet des milliers d'utilisateurs se retrouvent sur Internet et jouent un rôle de contributeur en ligne. Serge Proulx¹³ traite la contribution en ligne en tant que forme sociale. Selon lui les nouvelles pratiques de communication et d'échange sont marquées par les

¹² BOUYE, Édouard. Le Web collaboratif dans les services d'archives publics : un pari sur l'intelligence et la motivation des publics. *La Gazette des archives* [en ligne]. 2012, n°227, pp. 125-136. Disponible sur : <http://www.persee.fr/doc/gazar_0016-5522_2012_num_227_3_4974> (consulté le 02/01/2018)

¹³ Proulx, 2014, *Opt. Cit.* p. 18

idéaux d'une culture participative et orientées vers un imaginaire de la contribution. En effet on constate que depuis quelques années cette culture participative s'est installée peu à peu dans la société contemporaine. Ce phénomène est surtout marqué par les activités en ligne. Entre autres activités sur le Web : l'annotation (d'un document), l'indexation (*tagging*), la collaboration à la réalisation de documents collectifs la rédaction (autorship), l'organisation interne d'un document (structuration), la création d'hyperliens, l'organisation externe d'un ensemble de documents, le remixage de contenus, les commentaires à des blogs, la publication en ligne (post) : texte, image, son, vidéo, hyperlien, logiciel [Proulx, 2011]¹⁴. Toutes ces activités ont lieu sur le web social qui est devenu un espace participatif où chacun apporte sa contribution en publiant des vidéos sur Youtube ; en partageant une photo, un document sur Facebook ; en commentant une publication LinkedIn ; en contribuant à l'encyclopédie libre Wikipédia ; en relayant des informations sur Twitter ; etc.

Toutes ces pratiques d'information et de communication interpellent des questions économiques. Dans ce cas il importe beaucoup de soulever le problème de la transaction marchande. Le contributeur qui publie un contenu « *user generated content* » sur une plateforme de son choix, peut le faire par simple plaisir. L'internaute peut aussi venir chercher une utilité économique par ce geste contributif. Parfois ce geste peut être commercialisé par les propriétaires des plateformes de publication à l'insu du contributeur. Plus encore, ce bien informationnel peut être une monnaie d'échange entre les géants du Web. Dans cette perspective Serge Proulx affirme :

« Le contributeur est non seulement un « producteur de contenu » (*content producer*), il agit aussi - et parfois à son insu -, [...] comme « fournisseur de données » (*data provider*). Ainsi, par exemple, lorsqu'il publie un hyperlien ou une image dans son profil Facebook, il produit un contenu propre dans l'univers Internet et participe ainsi à l'accroissement du flux informationnel qui constitue le matériau primordial des sites en ligne. Simultanément il communique l'entreprise de la plateforme des métadonnées commercialement pertinentes [...] et qui sont à la base du processus de création de valeur dans le régime de l'économie de la contribution. »

L'internaute communique des informations de connexion qui peuvent être pertinentes pour les sites de commerce et les annonceurs par exemple. Ces contributions minuscules des petits usagers selon Proulx sont la source de la création de la valeur économique des géants. Proulx qualifie ces contributions de « passives non conscientes » et non intentionnelles. Elles se font

¹⁴ Proulx, Serge. La puissance d'agir d'une culture de la contribution face à l'emprise d'un capitalisme informationnel. Premières réflexions. *Journal du Mauss* [en ligne], 2011. Disponible sur : <<http://www.journaldumauss.net/?La-puissance-d-agir-d-une-culture>> (consulté le 09/03/2018)

souvent à l'insu des internautes. Pour expliquer cela, citons l'exemple donné par Proulx sur l'achat en ligne. Lorsque l'internaute effectue un achat sur Amazon, s'exprime sur Facebook ou procède à une recherche sur Google, une suite de données précises et commercialement pertinentes est fournie aux plateformes par les usagers à l'insu même de ces derniers, du simple fait de leurs activités en ligne. Leurs choix et leurs préférences, leurs goûts, les informations (parfois personnelles) laissées sur d'autres sites liés, sont automatiquement associées à leur adresse IP et envoyées dans des bases relationnelles servant à produire, par exemple, des systèmes de recommandations envoyées automatiquement à des acheteurs potentiels, ou des profils sophistiqués de consommation à la base des stratégies publicitaires ciblées visant à susciter une propagation virale [Proulx, 2014 p. 19].

Par ailleurs ces gestes contributifs peuvent être motivés, comme nous l'avons soulevé précédemment, par le plaisir. Évidemment la plupart des internautes ne sont pas motivés par l'appât du gain. La contribution en ligne peut aussi être justifiée par un désir de légitimation dans un groupe ouvert ou fermé. Cette gratitude recherchée entre confrères explique le plus souvent certaines activités contributives. Cela vient d'un besoin d'une certaine renommée de la part des internautes. Serge Proulx propose d'articuler ces gestes contributifs avec les logiques de don. Ainsi, il avance :

« Ces activités contributives s'ancrent plutôt dans le désir d'une réciprocité dans les échanges et, s'il y a lieu, elles peuvent faire l'objet d'une rétribution symbolique s'exprimant alors surtout en termes de réputation auprès des publics plus ou moins proches, ou de reconnaissance par les pairs ».

Concernant les significations attribuées à ces pratiques par les contributeurs eux-mêmes, il convient d'abord de remarquer que la plupart des cas, les contributeurs ne cherchent pas a priori de compensation financière pour leur geste contributif. Bien au contraire, un échange d'argent signifierait une marchandisation de leur contribution, ce qui aurait pour conséquence de détruire l'esprit de gratuité propre au climat de coopération et de confiance réciproque, et le « sentiment de partage » qui caractérisent ces réseaux et ces communautés d'échange d'informations et de services [Proulx, 2014, p. 19].

Cette contribution, présente dans l'univers numérique, a fait émerger dans une forme nouvelle un modèle classique de faire travailler la foule ; que l'on pourrait appeler le *crowdsourcing*. Le *crowdsourcing* est donc une forme de contribution reposant sur l'intelligence des foules. En effet ce type de pratique participative puise sa force dans le nombre important d'internautes et enregistre un succès considérable dans le monde des archives. Ainsi, avant d'aborder la pratique

du *crowdsourcing* dans les services archives et comment il contribue à valoriser les archives, nous allons d'abord le décortiquer. Nous allons donc étudier cette notion de *crowdsourcing* dans sa forme littérale, ensuite remonter à ses origines historiques et conceptuelles avant de terminer sur ses différents types.

C. Le *crowdsourcing*

C'est un concept anglais utilisé pour désigner une forme d'intelligence collective. Dans ce mémoire nous le considérons comme tel et nous essaierons de montrer ses pratiques au niveau des services d'archives et son rapport aux stratégies de valorisation des fonds d'archives. Mais avant tout cela, voyons d'abord ce qu'en disent les auteurs.

1. Définitions diverses, sens convergents

Le *crowdsourcing* signifie :

« La mobilisation de la foule comme ressource pour effectuer un travail mais il existe différentes définitions suivant que l'on prenne en compte ou pas la contractualisation, l'implication consciente, le volontariat, la collaboration... » [Szoniecky, 2017, p. 23].

Un mot qui a été défini plus d'une fois dans la littérature scientifique, le *crowdsourcing* trouve sa filiation conceptuelle dans des idéologies socialiste, marxiste, anarchiste, humaniste ou libéraliste. Ses origines historiques peuvent remonter très loin dans le temps avec les appels lancés au XVIIIe siècle par le gouvernement anglais aux scientifiques pour déterminer la longitude en mer d'un bateau, selon Andro et Saleh¹⁵.

Dans le Journal Officiel de la République Française, le *crowdsourcing* est défini comme étant de la production participative. Plus précisément c'est l'utilisation de la créativité, de l'intelligence et du savoir-faire d'un grand nombre de personnes, en sous-traitance, pour réaliser certaines tâches traditionnellement effectuées par un employé ou un entrepreneur.¹⁶

Reposant sur l'intelligence des foules et l'aspect collaboratif du Web, le *crowdsourcing* est littéralement défini comme de l'externalisation (*outsourcing* auprès de foules (*crowd*))

¹⁵ Andro, Mathieu, Saleh, Imad in Intelligence collective et archives numériques, *Opt. Cit.*, pp. 135-158

¹⁶ Production participative. *JORF*. 5 août 2014, n°0179, texte n° 91. Légifrance [en ligne]. Disponible sur : <https://www.legifrance.gouv.fr/affichTexte.do?jsessionid=3E3F01A4E2A178BD0687960539544F0F.tplgfr25s3?cidTexte=JORFTEXT000029331922&dateTexte=&oldAction=rechJO&categorieLien=id&idJO=JORFCON000029330499> (Consulté le 10 décembre 2017)

d'internautes selon l'expression de Jeff Howe proposée dans la revue *Wired Magazine* en juin 2006¹⁷.

Selon une définition publiée dans le deuxième numéro du *Journal of Information Science* en 2012, le *crowdsourcing* est :

« Un type d'activité en ligne pour laquelle un individu, une institution, une organisation à but non lucratif ou une société propose à un groupe hétérogène et indéfini d'individus de compétences variables, à travers un appel ouvert, le volontariat, la sous-traitance d'une tâche ou de micro tâches. L'externalisation de cette tâche, et pour laquelle, une foule d'internautes pourraient participer, apportant leur travail, leur argent, leurs connaissances et/ou expériences bénéficie toujours mutuellement à tous les associés. Les usagers recevront la satisfaction d'un type donné de besoins, qu'ils soient économiques, basés sur la reconnaissance sociale, l'estime de soi, ou le développement de compétences individuelles. » [EST 12]¹⁸

2. Participation involontaire des internautes

Contrairement aux auteurs précédents, Andro et Saleh attestent que le *crowdsourcing* peut aussi exister sous la forme d'une participation qui n'est pas obligatoirement et strictement volontaire. Les internautes peuvent bel et bien faire du *crowdsourcing* sans s'en rendre compte. C'est le cas de projets où les internautes contribuent en jouant et qui sont qualifiés de *gamification*¹⁹ ou *ludification*. Il peut aussi faire appel à la participation involontaire ou inconsciente des internautes, comme par exemple le projet reCAPTCHA. C'est un mécanisme de sécurité qui demande à l'internaute d'effectuer certains calculs, afin de prouver qu'il n'est pas un robot. Le système part du principe que si les utilisateurs résolvent correctement le CAPTCHA habituel, alors ils ont aussi déchiffré correctement le mot inconnu. Néanmoins, un mot n'est considéré comme vraiment reconnu que si plusieurs utilisateurs l'ont vérifié en obtenant le même résultat.²⁰

Autre exemple de participation involontaire : les millions de livres numérisés par Google Books et *océrisés*. Les mots n'ayant pas d'occurrences dans des dictionnaires sont ensuite renvoyés vers les internautes qui, pour des raisons de sécurité, sont contraints de ressaisir des mots parfois

¹⁷ BOUHAI, Nasreddine, SZONIECKY, Samuel, dir. *Archives numériques et intelligence collective : vers des écosystèmes de connaissances*. Londres : ISTE Editions, 2017, 283 p.

¹⁸ ESTELLES-Arolas E., GONZALEZ-Ladron-De-Guavara F., « Towards an integrated crowdsourcing definition », *Journal of Information Science*, vol. 38, n°2, 2012.

¹⁹ C'est transposer les mécaniques du jeu dans un domaine non-ludique pour résoudre des problèmes de la vie réelle ou améliorer une offre. El gamificator [en ligne]. Disponible sur : <http://www.elgamificator.com/gamification/gamification-definitio> (consulté le 09/03/2018)

²⁰ <https://developers.google.com/recaptcha/> (consulté le 11/03/2018)

déformés pour prouver qu'ils ne sont pas des robots. Ce faisant, en créant leurs comptes sur des sites web, ils contribuent involontairement à la correction de l'OCR des programmes Google Books et Google Maps ; et des archives du *New York Times*. [Andro et Saleh, 2017, p. 136]. Ce qu'il qualifie d'ailleurs de *crowdsourcing* implicite.

3. Origines historiques et conceptuelles du *crowdsourcing*

Bien qu'il soit un phénomène ancien, le *crowdsourcing* s'est beaucoup plus développé avec l'émergence des technologies de l'information et de la communication, en particulier avec le Web 2.0 qui favorise la liaison de plusieurs individus disséminés à travers le monde [Cardon, 2006]²¹. Ainsi, pour cerner ce phénomène et mieux l'appréhender, il convient de remonter à ses origines historiques, mais aussi conceptuelles.

a. Ses origines historiques

C'est un modèle économique qui trouve sa filiation dans les appels au peuple lancés par les gouvernements pour résoudre des problèmes scientifiques contre des récompenses dès le XVIII^e siècle. Les concours et les souscriptions publiques ; le libre-service et le libre accès ont permis au consommateur de prendre en charge une partie du travail du producteur jusqu'au modèle du « à la demande » qui lui a permis de prendre en charge la décision de mise en production elle-même [Andro et Saleh, 2017, p. 137].

Effectivement les premières pratiques participatives ont lieu au XVIII^e siècle. Le gouvernement anglais lança un appel aux scientifiques en 1714 afin qu'ils trouvent une solution pour déterminer la longitude en mer d'un bateau. Cela pourrait être à l'origine du *crowdsourcing* parce que faisant appel à l'intelligence d'un groupe pour trouver une solution à un problème.

Le tableau ci-dessous est tiré du chapitre d'ouvrage rédigé par de Andro et Saleh. Il récapitule toutes les dates importantes correspondantes à des événements significatifs, fondateurs du *crowdsourcing* et des sciences citoyennes.

1714	Le gouvernement anglais lance un appel aux scientifiques afin qu'ils trouvent une solution pour déterminer la longitude en mer d'un bateau. John Harrison, un charpentier et horloger remporte les 2000 livres de récompense parmi plus d'une centaine de concurrents dont Cassini, Huygens, Halley et Newton.
1726	Une ordonnance du Roi de France, Louis XV, demande aux capitaines de navires de rapporter des plantes et des graines des pays étrangers qu'ils visitent.

²¹ Cardon Dominique, Perrier Valérie Jeanne, Le Cam Florence *et al.*, « Présentation », *Réseaux* [en ligne], 2006/3 (n° 137), p. 9-25. Disponible sur : <https://www.cairn.info/revue-reseaux1-2006-3-page-9.htm> (consulté le 18/05/2018)

1750	L'astronome britannique Nevil Maskelyne calcule la position de la Lune pour la navigation en mer grâce à la confrontation des calculs de deux astronomes ayant effectué les calculs deux fois chacun puis ayant été vérifiés par un tiers.
1794	L'ingénieur français Gaspard de Prony organisa des micro-tâches d'additions et de soustractions auprès de quatre-vingts coiffeurs au chômage afin de développer des tables logarithmiques et trigonométriques détaillées.
1810	Nicolas Appert reçut 12 000 francs du gouvernement français suite à un appel à contributions pour ses nouvelles méthodes de conservation des aliments qui aboutiront aux boîtes de conserve.
1852	Le magasin « Au bon marché » est le premier magasin en libre-service. Désormais, les consommateurs accèdent directement aux marchandises sans passer par l'intermédiaire du marchand et prennent ainsi en charge une partie du travail du producteur.
1857	L'Oxford English Dictionary bénéficie, suite à un appel à contributions bénévoles, de plus de six millions de documents contenant des propositions de mots et des citations d'usages.
1884	La statue de la Liberté est financée par une souscription publique.
1893	Francis Dalton constate, à l'occasion d'un concours lancé sur un marché de bétail et pour lequel il s'agissait de deviner le poids d'un bœuf, que la moyenne des estimations d'une foule plus proche de la vérité que l'estimation des experts, laissant supposer l'existence d'une « sagesse des foules ».
1895	Le bibliothécaire James Duff Brown invente le libre accès en bibliothèque. Les lecteurs de la bibliothèque publique de Clerkenwell ont désormais accès directement à une partie des collections
18??	Dans le domaine de l'édition, les souscriptions publiques se multiplient afin de financer la publication de livres.
1900	La National Audubon Society (USA et Canda) organise un décompte annuel des oiseaux, le « Christmas bird count ».
1936	Toyota rassemble 27 000 personnes et choisit le meilleur dessin proposé comme logo de sa marque.
1938	Aux Etats-Unis, le Mathematical Tables Project mobilise 450 chômeurs victimes de la dépression économique et pilotés par un groupe de mathématiciens et de physiciens afin de calculer les tables de fonctions mathématiques, bien avant l'invention de l'ordinateur.
195?	L'ingénieur industriel de Toyota, Taiichi Ōno, invente le modèle du « juste à temps » ancêtre du modèle « à la demande » qui permettait de produire sans stocks ni invendus, en flux tendus, en fonction de la demande. Il s'agit, d'une certaine manière, d'externaliser auprès du consommateur la décision de mise en production elle-même. Ce modèle est à l'origine de la numérisation à la demande par <i>crowdfunding</i> et de l'impression à la demande.
1954	Le premier Téléthon aux Etats-Unis permet de recueillir un financement pour lutter contre l'infirmité motrice cérébrale.
1955	L'Opéra de Sydney est conçu et construit à la suite d'un concours public, qui encouragea les gens ordinaires de 32 pays à contribuer à ce projet de conception.
1979	Le sondage Zagat (un guide de restaurants) fonda ses évaluations sur une grande quantité de testeurs. Le projet fut racheté en septembre 2011 par Google.
1981	Le guide de voyage Lonely Planet est rédigé, pour sa troisième édition, de manière participative par des voyageurs indépendants.
1996	Naissance d'Internet Archive.

1997	Le livre à la carte : reproduction par fac-similés des livres conservés dans les bibliothèques (numérisation et impression à la demande)
1998	L'annuaire Dmoz, propose un contenu généré par ses utilisateurs. Le web 2.0 était né.
2000	Les plateformes philanthropiques de <i>crowdfunding</i> justgiving.com et de financements participatifs d'artistes artistshare.com voient le jour. Elles seront suivies par des multiples initiatives jusqu'à aujourd'hui.
2000	Distributed Proofreader : premier projet de transcription participative de livres.
2001	Naissance de Wikipédia.
2003	ESP Game : un jeu pour l'indexation des images.
2005	Amazon lance la plateforme de <i>crowdsourcing</i> <i>Amazon Mechanical Turk Marketplace</i> pour ses propres besoins et permet aussi de mettre en relation des sociétés et des institutions à la recherche et des travailleurs sur le web autour de micro-ondes.
2006	Espresso Book Machine pour l'impression à la demande <i>in situ</i> .
2006	Jeff Howe propose le terme de <i>crowdsourcing</i> dans la revue <i>Wired Magazine</i> en juin 2006.
2007	Google Books utilise reCAPTCHA pour faire corriger son OCR brut par les internautes.
2008	Le projet de <i>gamification</i> Fold.it permet de progresser, grâce à des jeux de puzzle, dans la connaissance des protéines.
2011	Le projet Good Judgement Project mobilise la sagesse de foules d'internautes à travers leurs prévisions géopolitiques qui rivalisent avec celles des experts du renseignement.
2011	Digitalkoot pour la correction de l'OCR sous la forme de jeux d'arcade.
2013	Le jeu vidéo <i>Star Citizen</i> récolte la somme de 30 044 586 \$.

Figure 1 : Origines historiques et chronologique du *crowdsourcing*²²

b. Ses origines conceptuelles

Le terme de *crowdsourcing* a été proposé pour la première fois par Jeff Howe dans la revue *Wired Magazine* en juin 2006. Il trouve sa filiation conceptuelle dans des idéologies aussi diamétralement opposées que les idéologies socialistes, libertaires, humanistes ou libérales selon Mathieu Andro et Imad Saleh. Pour eux l'idéologie californienne réaliserait la synthèse la plus favorable au développement du *crowdsourcing*. [Andro, Saleh, 2017, pp. 140-141]

Les idées fondatrices du *crowdsourcing* sont synthétisées ici dans des cartes heuristiques réalisées à partir des analyses de Mathieu Andro et Imad Saleh. Chacune des cartes représente une ou des idéologies spécifiques.

²²Mathieu ANDRO, Imad SALEH, 2017, *Opt. Cit.*, pp. 137-139

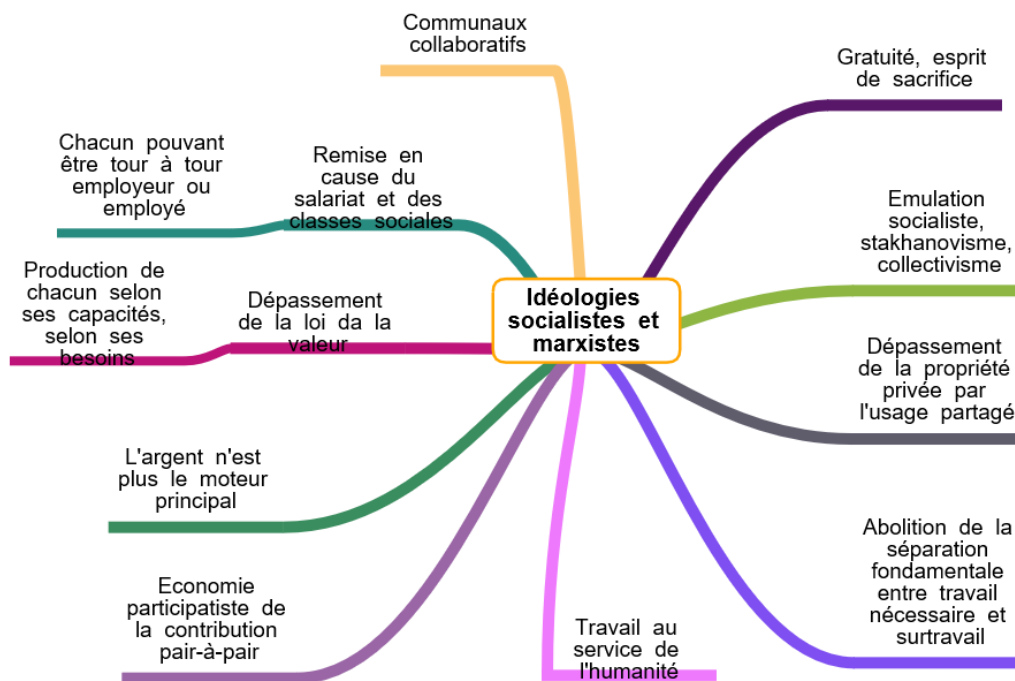


Figure 2 : Représentation des idéologies marxistes et socialistes

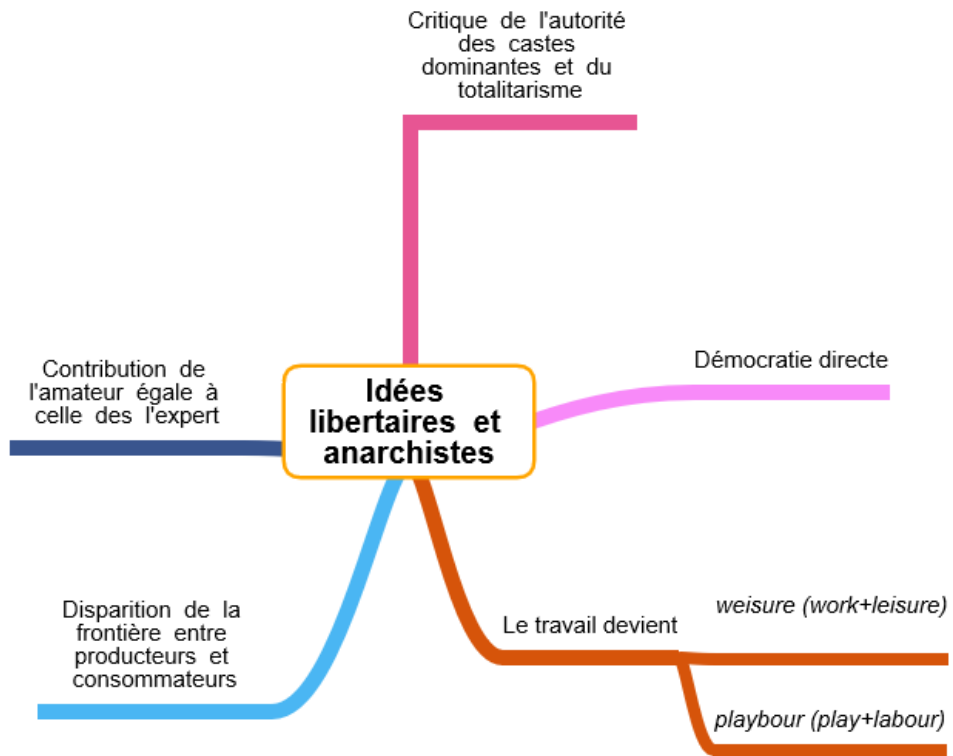


Figure 3 : Représentation des idées libertaires et anarchistes

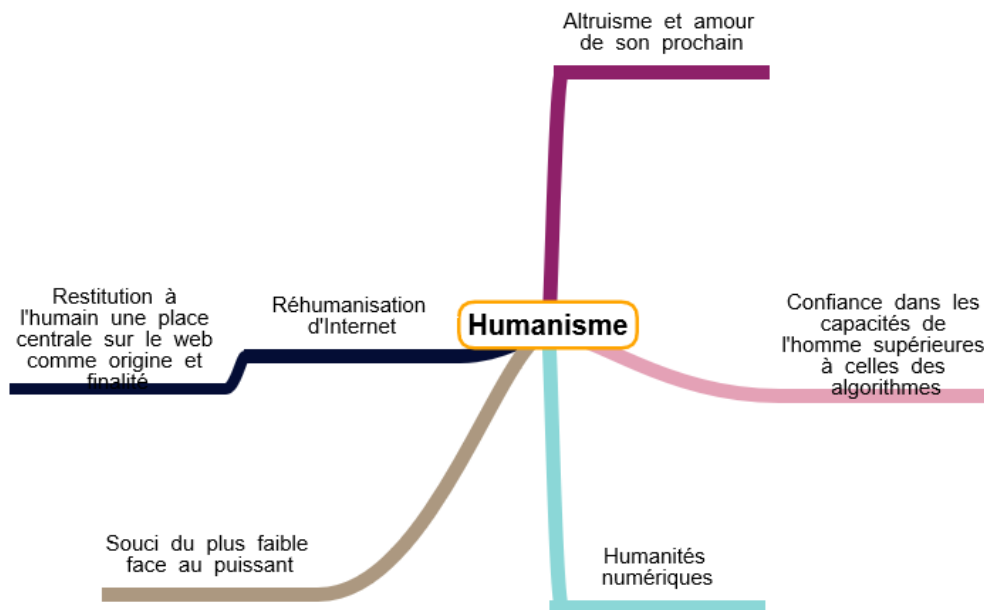


Figure 4 : Représentation des idées humanistes

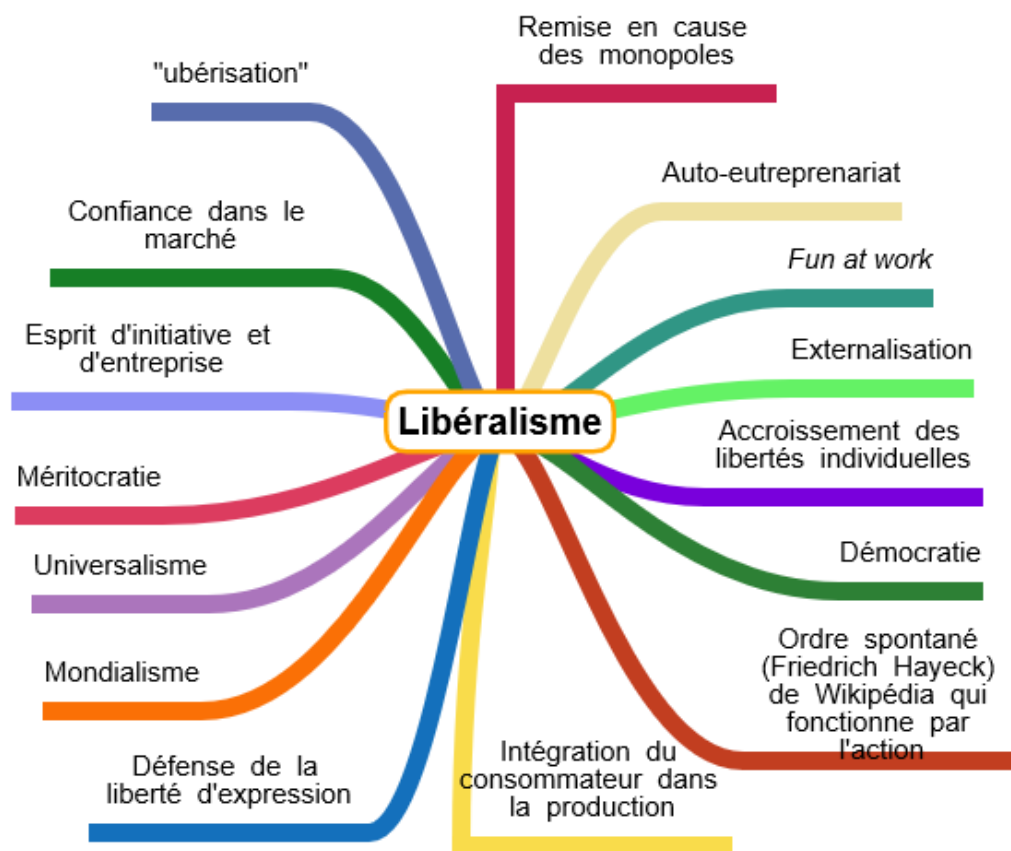


Figure 5 : Représentation du libéralisme

4. Typologie du *crowdsourcing*

Du point de vue quantitatif, les auteurs distinguent deux types de *crowdsourcing*. Les spécialistes de la question ont en effet identifié une typologie des projets de *crowdsourcing* en fonction du degré d'engagement du public. Il existe le *crowdsourcing* participatif ou contributif où la participation du public est très limitée à des micro-tâches qui demandent peu d'investissement. En effet ce sont les institutions qui conçoivent et pilotent ces projets, et les internautes s'en tiennent qu'à la production de contenu. Par ailleurs le *crowdsourcing* collaboratif permet aux internautes d'interagir entre eux et de s'investir fortement avec la co-création de contenu. Ainsi, les internautes pourront influencer la politique, le périmètre ainsi que les objectifs du projet ; et mieux, ils pourront même être à l'origine de ces projets.

Au-delà de cette distinction quantitative, Mathieu ANDRO et Imad SALEH proposent une typologie plus qualitative. Ils distinguent en premier le *crowdsourcing* explicite du *crowdsourcing* implicite. Avec ce dernier, les institutions font recours au travail involontaire des internautes en leur faisant corriger l'OCR avec reCAPTCHA de Google par exemple. Dans ce cas la plupart des internautes ne sont même pas au courant de leur participation. Ils ne sont pas conscients que les termes qu'ils saisissent sont utilisés par une entreprise ou pour alimenter une base de données. Le *crowdsourcing* explicite, au contraire, suppose un recours au travail volontaire des internautes. Les internautes sont très conscients de leur participation et sont mêmes impliqués dans les projets. À partir de là les auteurs établissent deux catégories : le *crowdsourcing* bénévole et le *crowdsourcing* rémunéré. De plus il y a le recours au travail des internautes sous la forme de jeux, autrement appelé la *gamification*, qui peut aussi rentrer dans la catégorie de *crowdsourcing* implicite. C'est le cas par exemple de DigiTalkoot, un projet lancé en 2011 par la bibliothèque nationale de Finlande et Microtask, et qui consiste à la reconnaissance de texte numérisés en jouant à des jeux [Ando, Saleh, 2017, p. 143].

D. Autres notions liées au *crowdsourcing*

Dans la littérature scientifique on trouve plusieurs notions liées directement ou indirectement à l'intelligence collective et au *crowdsourcing*. Cette nouvelle pratique participative, reposant sur le Web et qui place l'utilisateur au centre des activités sur l'Internet, est favorisée par la deuxième génération du Web qu'on nomme le « Web 2.0 ». Ces termes rattachés au *crowdsourcing* et à l'intelligence collective, caractérisent souvent les projets collaboratifs. Ils peuvent être liés soit à l'environnement Internet soit aux pratiques elles-mêmes.

1. Le Web 2.0

Les pratiques du *crowdsourcing*, la contribution en ligne, l'indexation sociale, sont toutes facilitées par ce qu'on appelle le « Web 2.0 ».

C'est une expression qui se réfère à une nouvelle génération de développement sur le web et de « web design ». Elle fut inventée par Darcy Di Nucci en 1999 [Lucien, Quoniam, 2010, p. 10]. En octobre 2004 lors d'une conférence brainstorming entre la société O'Reilly Média (du nom de son fondateur Tim O'Reilly) et la société MediaLive International, Dale Dougherty, membre d'O'Reilly, suggéra que, loin de s'être effondré suite à l'explosion de la bulle internet en 2001, le Web n'avait jamais semblé aussi important et novateur. Tim O'Reilly lia cette appellation à l'idée de changements cumulatifs dans le développement du web ainsi qu'aux nouveaux usages

sur la toile, mis en place par les utilisateurs finaux. Une seconde conférence fut organisée un an plus tard en octobre 2005, enrichie de la réflexion que Tim O'Reilly développe dans un article désormais considéré comme fondateur : « What is web 2.0 ? ». L'engouement, que n'avait pas provoqué la première session, devient général et la conférence se voit une nouvelle fois reconduite et renommée en « sommet du Web 2.0 » en novembre 2006. [Chaimbault, 2007, p. 5].

Cette nouvelle génération du Web repose surtout sur un renversement de la logique *top-down* du web initial. Ainsi, Chaimbault avance :

« Alors que le Web 1.0 « descendait » vers l'utilisateur pour lui proposer contenus et services, le Web 2.0 met l'accent sur une nouvelle forme d'interactivité qui place l'utilisateur au centre de l'internet et se veut plus social et collaboratif [Chaimbault, 2007, p. 6].

Selon Quoniam et Lucien²³ :

« La première génération du Web reproduisait un modèle de communication dit « *one to many* » commun aux médias traditionnels alors que les dispositifs socio-techniques 2.0 proposent de nouveaux usages reposant sur un modèle de communication « *many to many* ». Le web 2.0 traduit alors le passage de l'interactivité à l'interaction et contribue ainsi à la construction de réseaux qui ne se basent plus sur l'échange d'informations mais sur le partage du savoir. »

Il favorise aussi l'apparition d'un contenu d'information élaboré par les utilisateurs finaux, plutôt que par les « producteurs finaux ».

Le « social bookmarking » remet au cœur du débat le fait que le « nouveau pouvoir » réside dans le partage de l'information et non dans la détention d'information, sans oublier le rôle joué par ce « regard des utilisateurs finaux » pour des approches sérendipistes, indispensables pour l'innovation, moteur de développement économique.²⁴

Pour ces auteurs il s'agit d'un ensemble de principes et de pratiques qui permettent au Web de devenir une plateforme d'échange entre les utilisateurs, les services et les applications en ligne plutôt qu'une vitrine. C'est la place de l'internaute qui est bouleversée, celui-ci devenant le principal diffuseur d'informations à travers les *blogs*, *wiki*, *tags* et réseaux sociaux. C'est une véritable organisation du web par les internautes.

²³ LUCIEN, Arnaud, QUONIAM, Luc in *Intelligence collective et organisation des connaissances*, *Opt. Cit.*, p. 15

²⁴ Quoniam, *Opt. Cit.*, p. 11

Cette nouvelle organisation du Web favorise l'émergence de nouvelles pratiques qui permettent de décrire avec les mots de l'utilisateur, de chercher et de retrouver l'information de manière plus aisée.

2. L'indexation sociale (métadonnées sociales) ou *folksonomies* ou indexation communautaire

L'ensemble des tags assignés par les internautes constitue une *folksonomie*, contraction de *folk* (les gens, le peuple) et *taxonomy* (classification hiérarchisée). On distingue deux types de *folksonomies*, les « étroites » (*narrow folksonomies*) et les « générales » (*broad folksonomies*) [Durieux, 2010, p. 71].

La construction des *folksonomies* étroites (représentées par des systèmes tels que Flickr repose sur l'indexation par l'utilisateur de ses propres ressources uniquement, tandis que les *folksonomies* générales sont le résultat de l'attribution de tags par les internautes à leurs propres ressources, mais également à celles d'autrui. Cette distinction est essentielle car elle démontre que toute *folksonomie* ne revêt pas forcément d'aspect collectif et collaboratif. Pour bien marquer cette différence, certains auteurs préfèrent d'ailleurs le terme de *personomie* à celui de *folksonomie* étroite [Le Deuff, 2006]. « En restreignant le nombre de tagueurs par ressource, les *folksonomies* étroites ne bénéficient pas de la richesse que peut apporter la masse. En effet, l'attribution de tags par plusieurs internautes à une même ressource multiplie considérablement le nombre de chemins d'accès, augmentant de cette manière ses chances d'être trouvée via une recherche par mot-clé. » [Durieux, 2010, p. 71]. Ce n'est plus alors un seul professionnel qui indexe mais toute une communauté d'utilisateurs.²⁵

Les *folksonomies* sont, en quelque sorte, un exemple de *crowdsourcing*. Olivier Le Deuff en donne une définition centrée sur sa différence avec les classifications normalisées et le libre choix des mots laissé aux publics :

« Les *folksonomies* constituent la possibilité pour l'utilisateur d'indexer des documents afin qu'il puisse plus aisément les retrouver grâce à un système de mots-clés. Le concept est lié à l'accroissement et à l'accélération de la production d'informations. Le système diffère donc nettement en cela des systèmes classificatoires classiques, comme la classification décimale universelle ou la classification Dewey, qui s'inscrivent dans des processus plus longs et dont le but est d'obtenir un classement cohérent de documents physiques dont le contenu est inscrit dans la durée. Les *folksonomies*, au contraire, ne reposent sur aucun thésaurus, ce qui confère à l'utilisateur une liberté totale quant au choix des mots-clés. Les *folksonomies* sont donc

²⁵ DURIEUX, Valérie « Collaborative tagging et folksonomies. L'organisation du web par les internautes », *Les Cahiers du numérique* 2010/1 (Vol. 6), p. 69-80.

initialement centrées sur l'utilisateur. Ce dernier les utilise dans un but personnel, afin d'organiser son propre système d'information. » [Le Deuff, 2006, p. 66].

Le terme qui revient le plus souvent dans les *folksonomies* est celui de tag qui désigne en fait un mot-clé, une catégorie ou une métadonnée. Le mot tag signifiant en anglais : étiquette de balisage, étiquetage, fléchage, marquage, voire traçage, il est possible de rencontrer des traductions comme *tagage collaboratif*.

Mais qu'est-ce qu'un tag ?

Voici la définition qu'en donnent Guy Marieke et Emma Tonkin :

« Une définition simple serait de dire que les tags sont des mots-clés, des catégories de noms, ou des métadonnées. Essentiellement, un tag est simplement un jeu de mots-clés librement choisi. Cependant, du fait que les tags ne sont pas créés par des spécialistes de l'information, ils ne suivent aucune indication formelle. Cela signifie que ces items peuvent être catégorisés avec n'importe quel mot définissant une relation entre la ressource en ligne et un concept issu de l'esprit de l'utilisateur. Un nombre infini de mots peut être choisi, dont quelques-uns sont issus de représentations évidentes tandis que d'autres ont peu de signification en dehors du contexte de l'auteur du tag. » [Marieke et Tonkin, 2004].

3. Les « archives participatives »

L'utilisation des techniques du Web 2.0 et de ce nouveau langage lié à la participation des utilisateurs, favorisent de l'émergence et l'utilisation de nouveaux concepts comme les « archives participatives ».

Selon Pauline Moirez les « archives participatives » constituent un mot-concept pour souligner l'investissement approfondi des usagers des archives. Ainsi, l'archiviste américaine Kate Theimer²⁶ définit les archives participatives comme étant :

« Un organisme, un site ou une collection auxquels des personnes qui ne sont pas des professionnels des archives apportent leur connaissance ou ajoutent des contenus, généralement dans un contexte numérique en ligne. Il en résulte une meilleure compréhension des documents d'archives. »

Les archives participatives s'appuient donc sur un véritable investissement des internautes, sur la mise en œuvre de compétences et connaissances de caractère plus ou moins scientifique. Il peut s'agir d'une contribution des usagers à la description des contenus (indexation collaborative, identification de photographies, etc.), voire d'une véritable co-construction

²⁶ THEIMER, K., *Exploring the Participatory Archives*, 2011, Disponible sur : <http://www.archivesnext.com/?p=2319> (consulté le 21-10-2017)

patrimoniale, via un apport de contenus scientifiques ou de matériaux patrimoniaux par les usagers [MOIREZ, 2012, p. 189].

Dans cette définition de l'archiviste américaine, l'aspect numérique des archives participatives est important à souligner. Ainsi, il faut comprendre que cette activité ne peut se faire que dans un environnement numérique. Celui-ci offre de multiples opportunités aux professionnels de l'information et aux archivistes en particulier. Autre aspect important est le fait que des non professionnels apportent leurs contributions à la description des documents. Ainsi les utilisateurs ont une meilleure compréhension des documents puisqu'ils les ont décrits avec leurs propres termes.

Le succès des archives participatives est considérable grâce à la masse des documents d'archives mais aussi à leur caractère unique qui constitue une spécificité par rapport aux autres types de fonds documentaires. Cette remarque importante sur la nature des fonds d'archives sera développée plus largement dans la troisième partie de ce mémoire.

Pour revenir à la notion d'archives participatives, elle désigne toutes les opérations de *crowdsourcing* mises en place pour l'identification, la description, l'indexation au niveau de la pièce et la transcription des contenus pour permettre une recherche en plein texte. En effet Pauline Moirez souligne que les archives conviennent très bien à un usage encadré et accompagné en salle de lecture car elles sont traditionnellement décrites à un niveau de granularité médian (au niveau de l'« article », c'est-à-dire une unité intellectuelle correspondant généralement à un dossier ou à un registre). Or ce n'est plus adapté aux usages actuels de la recherche du Web [MOIREZ, 2012, p. 189]. C'est dans un souci de visibilité sur le Web et d'accessibilité de l'information aux utilisateurs que des projets collaboratifs ont vu le jour dans le monde des archives un peu après les réalisations faites dans les bibliothèques et les musées. Car comme nous allons le détailler plus bas, les archivistes ont accusé un certain retard par rapport aux bibliothécaires et conservateurs de musée. Ce décalage peut être expliqué comme le précise Pauline Moirez :

« Cette problématique de la masse explique parfois un certain décalage avec les réalisations de médiation numérique et de présence sur le web social d'autres domaines culturels : la priorité des archivistes reste la numérisation et la mise en ligne des documents, là où les musées par exemple vont centrer leurs efforts sur la médiation et l'interaction avec les usagers même s'ils ne fournissent que très peu de contenus en ligne »

Les archives participatives peuvent aussi faire référence au Web participatif que les services d'archives utilisent pour collecter des documents d'archives privées et accroître leurs collections. Les techniques du Web participatif sont aussi utilisées pour l'identification de documents iconographiques, auxquels il est impossible d'accéder par un moteur de recherche s'ils ne disposent pas d'un minimum de données descriptives. [MOIREZ, 2012, p.189].

Ces différents types de projets mis en place par les services d'archives et qui peuvent être appelés sous le nom d' « archives participatives », ont été suscités par des motivations de la part des archivistes et des intérêts des publics. Ces projets, très diversifiés d'un service d'archives à un autre, diffèrent aussi d'un type de fonds à un autre. En effet dans les services d'archives, nous constatons une multitude de projets en fonction des stratégies et selon les objectifs visés. Cela nous amène donc à entamer la deuxième partie de ce mémoire afin d'explorer ce large panel de pratiques du *crowdsourcing* dans le domaine des archives.

II. CROWDSOURCING DANS LES SERVICES D'ARCHIVES

Depuis quelques années on a vu apparaître des projets de *crowdsourcing* dans beaucoup de services d'archives. Ces services d'archives par l'intermédiaire des archivistes visent plusieurs objectifs en initiant ces projets, notamment la mise en valeur des fonds d'archives. La question principale qui nous occupe dans ce mémoire est : « comment ces projets participent-ils à la valorisation des archives ? » Pour avoir des éléments de réponse à cette problématique nous avons mené une enquête auprès des services d'archives qui ont mis en place des projets de *crowdsourcing*. Dans les lignes qui suivent nous présentons l'enquête ainsi que les outils utilisés pour questionner les services d'archives. Nous donnerons ensuite quelques explications sur les critères de choix des projets que nous présenterons en détails. Enfin nous concluons sur une catégorisation que nous avons faite sur les différents types de projets que nous avons identifiés.

Dans son article publié en 2012, Édouard Bouyé²⁷ présente 30 services publics d'archives qui ont mis en place des projets participatifs. En mars 2018, sur le Portail National des Archives, 41 services d'archives sont identifiés, incluant les Archives Nationales. En plus de ceux-là, nous ajoutons le Ministère des Armées avec son projet d'indexation collaborative. Ainsi notre corpus est essentiellement constitué de ces services publics d'archives.

A. Enquête auprès des services d'archives ayant mis en place des projets de *crowdsourcing* : méthodologie adoptée

Pour cette enquête nous n'avons pas mené des entretiens, nous avons administré un questionnaire dont les détails sont décrits dans la partie suivante. Mener un entretien auprès des personnes concernées aurait été très enrichissant mais le temps et les moyens disposés ne nous l'ont pas permis dans ce travail. En plus c'était déjà difficile d'obtenir des réponses au questionnaire car les services d'archives n'étaient pas réactifs, alors ce n'était pas envisageable de faire des entretiens en plus. Même sur des aspects importants dans les réponses que nous avons voulu creuser, nous n'avons pas obtenu des éclaircissements.

Dans cette étude nous voulions aussi enquêter les contributeurs aux projets mais nous n'avons pas pu le faire. Nous avons contacté les Archives nationales pour avoir les coordonnées de quelques contributeurs afin de les questionner, mais la chargée du projet nous a expliqué que cela était confidentiel. Dans les clauses de confidentialité, les organisateurs du projet sont tenus

²⁷ *La Gazette des archives*, n°227/année 2012-3. Disponible sur : <http://www.persee.fr/doc/gazar_0016-5522_2012_num_227_3_4974> (consulté le 21/03/18)

de ne pas divulguer à un tiers les informations personnelles liées aux contributeurs. Néanmoins nous avons voulu essayer avec d'autres services d'archives mais par manque de temps et pour éviter une charge de travail supplémentaire, nous avons fini par abandonner cette partie de l'enquête. C'est aussi pour le cas des généalogistes. Nous avons identifié une plateforme de *crowdsourcing*, GeneaNet, mise en place par des associations de généalogistes où les internautes sont appelés à indexer, à identifier des documents d'archives, ajouter des cartes postales par exemple, etc. Ce type d'acteur est très intéressant à questionner car les points de vue sur les projets, les stratégies ainsi que les objectifs et les motivations peuvent être très différents de ceux des archivistes. Mais notre interlocuteur était un informaticien qui est en charge de la maintenance. Malheureusement c'est le contact unique présent sur le site et nous ne nous sommes pas bien compris. De ce fait nous n'avons pas pu avancer dans cette partie.

Pour résumer, nous avons donc administré un questionnaire aux services publics d'archives qui ont mis en place des projets collaboratifs. Les contacts de ces services ont été obtenus à partir du site *France Archives* qui permet de rechercher dans l'annuaire en ligne par simple saisie du nom du service d'archives en question.

1. Conception des outils d'enquête

Élaboration du questionnaire

Pour notre enquête auprès des services d'archives, nous avons élaboré un questionnaire. Le questionnaire a été conçu avec un outil en ligne, *Google Forms*. Les questions ont été élaborées dans un document texte et ont fait l'objet de beaucoup de modifications. Elles ont été formulées en partie à partir des questions de recherche et des objectifs visés dans l'étude. Les premières questions ont été un peu vagues mais avec l'appui de notre directrice de mémoire, nous avons pu reformuler et préciser des aspects qui ne l'étaient pas. Pour tester le questionnaire nous avons eu l'appui d'un archiviste très motivé et disponible. Après notre entretien téléphonique pour la présentation de l'étude, nous lui avons soumis le questionnaire par mail. Et c'est à partir de ses remarques et corrections que nous avons ajouté des questions sur des aspects essentiels et en reformulé d'autres. Son appui nous a été d'une très grande utilité car nous avons eu à la fois une vue d'ensemble sur le questionnaire et un regard critique sur certains points. Le questionnaire a été modifié au fur et mesure de nos lectures et des réponses qui ont été récoltées.

2. Présentation de la population et de l'échantillon de l'enquête

En matière de projets de *crowdsourcing* dans le domaine des archives, il existe de nombreux acteurs qui sont impliqués. Ils peuvent être soit des spécialistes du domaine c'est-à-dire des

archivistes (services d'archives) soit des non-professionnels comme c'est le cas des généalogistes (associations) soit les internautes qui participent aux projets (usagers).

a. La population : les services publics d'archives

Dans cette étude nous nous sommes intéressées au premier type d'acteurs pour constituer notre échantillon. Dès l'élaboration de l'enquête nous avons eu l'idée de questionner les généalogistes ainsi que les usagers qui participent aux projets mis en place par les archivistes. Mais comme nous l'avons expliqué précédemment cela n'a pas pu aboutir. La population enquêtée est donc constituée des services publics d'archives par l'intermédiaire des archivistes et / ou des responsables qui ont porté des projets de *crowdsourcing*.

b. L'échantillon : les services d'archives enquêtés

La majorité de l'échantillon est constitué par les services d'archives des départements car beaucoup d'initiatives de projets de *crowdsourcing* ont été prises par ces types de services d'archives. Ils sont au nombre total de 34 mais pour l'enquête nous en avons choisi 19. Ceux-ci ont été choisis de façon aléatoire suivant le type de projet qui nous intéresse, mais aussi selon l'accessibilité et la disponibilité de la documentation sur le projet. Les Archives Nationales ont aussi fait partie de l'échantillon puisqu'étant actuellement parmi les acteurs majeurs dans ce sens. En plus de ceux-ci, nous avons repéré les services d'archives des communes. Nous en avons identifié 6 qui ont mis au point des projets sur différents types de fonds. Nous avons aussi ajouté le Ministère des Armées qui a lancé un projet d'indexation collaborative de grande envergure depuis mars 2013.

Ainsi, le questionnaire a été administré aux 27 services d'archives dont leurs projets sont présentés ci-dessous et qui sont choisis selon des critères précis. L'envoi du questionnaire a été facilité par notre directrice de mémoire qui nous avait communiqué les coordonnées de quelques personnes susceptibles de répondre ou de nous rediriger vers d'autres personnes qui sont concernées par des projets de *crowdsourcing*. La plupart des coordonnées des services d'archives ont été trouvées dans « l'annuaire des services » sur le Portail National des Archives de France.

Le tableau suivant récapitule les services d'archives enquêtés ainsi que leurs projets. Dans le tableau suivant, nous marquons un « oui » dans les cases correspondantes aux services d'archives et institutions répondants dans la colonne « réponse au questionnaire » :

Numéro	Services d'archives	Projets	Réponse au questionnaire
1	Archives départementales de l'Ain	Projet d'Indexation des Registres paroissiaux et état civil	
2	Archives départementales d'Alpes Maritimes	Projet d'identification d'images et d'édition de textes anciens	
3	Archives départementales des Ardennes	Projet ... de registres des matricules	
4	Archives départementales de l'Aube	Projet d'un module d'annotation collaborative de l'état civil et recensement de la population	
5	Archives départementales de la Charente	Projet d'indexation de registres matricules	
6	Archives départementales et patrimoine du Cher	Projet d'annotation collaborative des registres paroissiaux et d'état civil	Oui
7	Archives départementales de la Haute-Garonne	Projet d'identification de photographies	
8	Archives départementales de la Haute-Loire	Projet d'indexation des fiches nominatives des registres matricules	Oui
9	Archives départementales de l'Hérault	Projet d'indexation collaborative des registres de matricules	Oui
10	Archives départementales d'Indre-et-Loire	Projet d'identification d'images de la <i>Base Collections de Touraine</i>	Oui
11	Archives départementales de la Lozère	Projet d'identification d'images	Oui
12	Archives départementales de la Marne	Projet d'indexation des registres matricules	
13	Archives départementales de l'Oise	Projet d'indexation collaborative des recensements de population	Oui
14	Archives départementales du Pas-de-Calais	Projet d'édition d'un carnet de soldats de la Grande guerre sur Wikisource	Oui
15	Archives départementales du Puy-de-Dôme	Projet d'indexation collaborative des registres matricules	
16	Archives départementales du Var	Projet d'annotation collaborative et d'identification de documents et de photographies	Oui
17	Archives départementales de Vendée	Projet de correction et d'enrichissement des	

		relevés nominatifs (registres paroissiaux, état civil et notaires) ; dictionnaires collaboratifs, l@boratoire des internautes, éphéméride vendéenne	
18	Archives départementales de l'Yonne	Projet d'indexation des matricules militaires	Oui
19	Archives départementales des Yvelines	Projet de transcription collaborative des testaments des Poilus	Oui
20	Archives Bordeaux Métropole	Projet d'identification d'images	Oui
21	Archives municipales de Béthune	Projet d'indexation collaborative des registres de l'état civil et les tables décennales	
22	Archives municipales et communautaires d'Orléans	Projet d'annotation collaborative des registres paroissiaux, de l'état civil, des recensements de population	Oui
23	Archives municipales de Saint-Brieuc	Projet d'annotation d'un fonds d'érudit	
24	Archives municipales de Saint-Denis	Projet d'identification d'images	
25	Archives municipales de Saint-Etienne	Projet d'indexation collaborative des registres d'état civil, les recensements et les listes électorales	Oui
26	Archives Nationales de France	Projet d'édition d'un carnet de guerre des Poilus	Oui
27	Ministère des Armées/ Secrétariat Général pour l'Administration / Direction des Patrimoines, de la Mémoire et des Archives / Délégation des Patrimoines Culturels / Bureau de la Politique des Archives et des Bibliothèques	Projet d'indexation collaborative (Mémoire des Hommes)	Oui

Figure 6 : Services d'archives et institutions enquêtés

3. Les services d'archives et institutions répondants

Sur un échantillon de 27 individus, nous avons récolté 15 réponses soit un taux de 53,5 %. Ce taux de réponses peut être jugé satisfaisant puisqu'il est bien au-dessus de la moyenne. Néanmoins il reste en dessous de nos attentes. Nous visions un taux de réponse de 80% pour

que notre échantillon soit le plus représentatif possible. Les services publics d'archives des départements et des communes sont identifiés ci-dessous sur la carte légendée. Nous avons choisi de les représenter ainsi car ils représentent 92% de l'échantillon, les 8% restant sont constitués par les Archives Nationales et le Ministère des Armées.

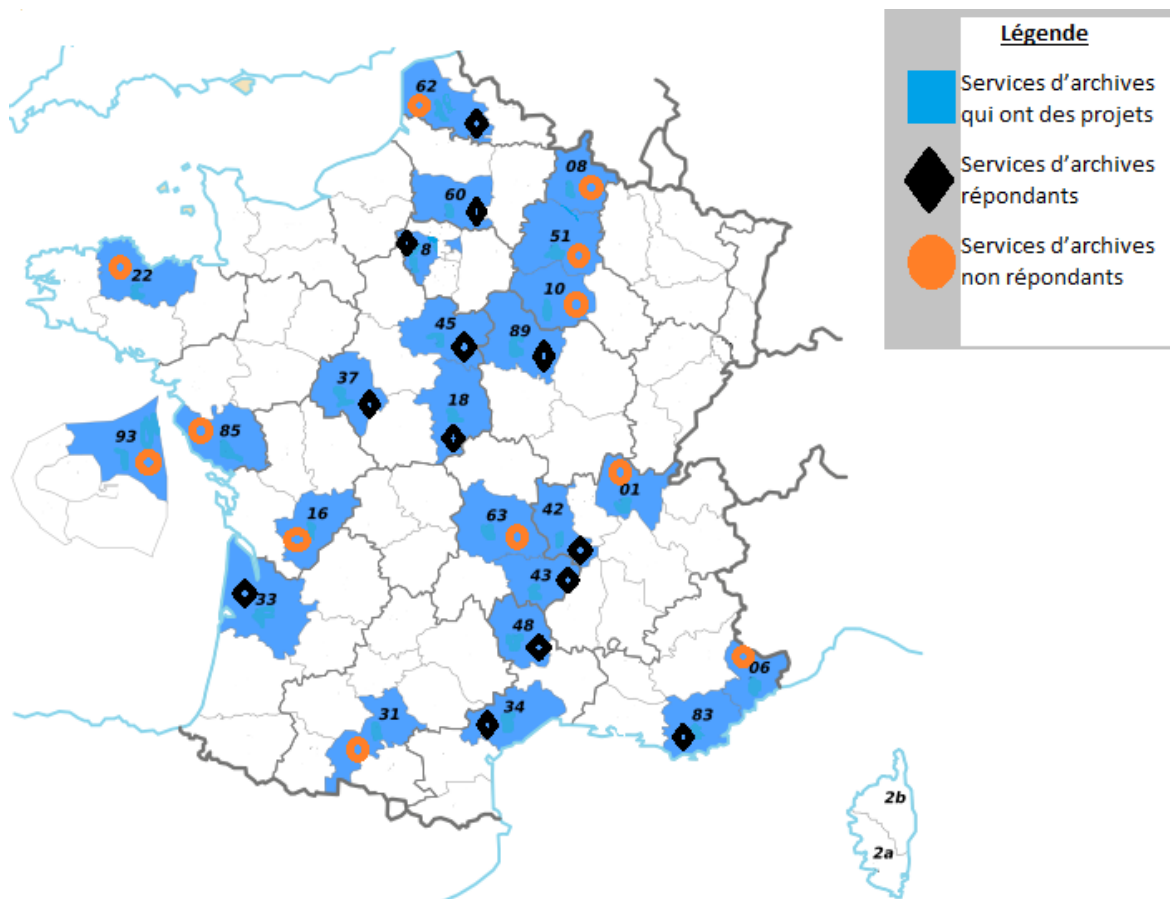


Figure 7 : Cartographie des services d'archives de l'échantillon

Sur cette carte nous avons représenté les 25 services d'archives départementaux et municipaux de notre échantillon qui ont mis en place des projets collaboratifs. Sur la carte ils sont répartis par départements. Les paramètres du site²⁸ ne nous ayant pas permis de représenter à la fois les départements et les communes, nous avons choisi de travailler sur la base de la carte des départements de France. Néanmoins retenons que les traits gris épais représentent les délimitations des régions et les traits gris fins les délimitations des départements. Les codes postaux sont aussi un indice d'identification des départements. Les formes ovale (orange) et losange (noir) indiquent le statut de la réponse expliqué dans la légende. Dans le cas du département du Pas-de-Calais nous avons les Archives départementales du Pas-de-Calais et les

²⁸ www.drawmeagraph.com

Archives municipales de Béthune. Nous avons donc dans ce département deux symboles : losange pour les Archives départementales du Pas-de-Calais qui ont répondu et ovale pour les Archives municipales de Béthune qui n'ont pas répondu.

a. Difficultés rencontrés

Le taux de réponses est moyen, sur 27 services enquêtés 15 réponses ont été récoltées soit 53,5%. Je peux dire que c'est relativement faible par rapport à l'objectif fixé. J'envisageais un taux largement au-dessus de la moyenne soit 80%. Car dans cette étude je visais une forte représentativité afin de pouvoir répondre à la problématique.

Comme toute enquête, les réalités c'est-à-dire les temps de réponses, m'ont rattrapées. L'obtention même des 15 réponses n'a pas été facile. Les services d'archives ont été relancés par mail trois fois pour certains et deux fois pour d'autres. Les relances par mail ont eu de l'impact sur les délais de réponses. En effet après les mails de relance, certains ont pris connaissance du questionnaire et d'autres ont répondu. Cependant il n'y avait pas beaucoup de réactivité de la part des services d'archives. J'ai donc recouru aux relances par téléphone. Les appels ont eu beaucoup plus d'effet que les mails, mais cela ne m'a pas permis d'atteindre notre objectif qui était un taux de réponse de 80%.

En plus j'ai eu des difficultés dans l'exploitation des réponses. Certaines personnes ont mal répondu à certaines questions, d'autres n'ont pas été précises. Il fallait donc aller une fois de plus visiter les sites Internet des services d'archives pour plus de précisions sur certaines réponses. Mais cela n'était possible que sur certaines réponses liées aux projets, par exemple : projet en cours ou terminé, utilisation du site web de l'institution ou d'une plateforme dédiée, etc. Sur certains points beaucoup plus complexes, il n'était pas possible d'aller vérifier. Je me suis donc contentée des réponses, même si elles sont parfois incomplètes. Dès l'élaboration de l'enquête, j'ai eu conscience du choix pour le questionnaire qui n'était pas adapté à certaines questions. C'est pour cela que je les avais reformulées le plus précisément possible. Dans tous les cas c'est mon choix et je l'assume car j'ai considéré que le traitement des réponses d'un questionnaire serait beaucoup plus rapide que celui d'un entretien. De plus, vu les délais et l'indisponibilité des personnes à répondre au questionnaire, il était difficile d'envisager des entretiens.

B. Panorama du *crowdsourcing* dans les services d'archives

Aujourd'hui, de plus en plus répandues sur les sites Internet des services d'archives, les pratiques collaboratives, en anglais *crowdsourcing*, permettent aux internautes de participer à la valorisation des archives et de faciliter les recherches en ajoutant de l'information (mots clés par exemple) aux archives numérisées et diffusées en ligne. Les pratiques collaboratives peuvent prendre plusieurs formes, de l'identification de photographies à la transcription de textes en passant par l'annotation et l'indexation collaboratives.

L'intérêt grandissant et le besoin d'améliorer la recherche dans les fonds concernés ont fait que ces services se sont beaucoup développés dans les archives. Au-delà de ces deux raisons, s'ajoute l'intégration de ces modules dans les logiciels disponibles sur le marché. Aujourd'hui il existe même des plateformes spécialement dédiées à ces types de pratiques.

Le choix de notre corpus dans le cadre de cette étude est fondé sur plusieurs critères. Nous avons choisi les projets qui ont un impact direct sur les fonds concernés. C'est-à-dire les projets qui utilisent l'intelligence et les connaissances des publics pour compléter, enrichir les fonds, ce qui participe à les valoriser dans la recherche.

Nous avons exclus de notre étude les projets de numérisation, bien vrai qu'ils soient très répandus dans les services d'archives. L'étude concerne donc les projets d'annotation, d'indexation, de transcription et d'édition de carnets, sans oublier l'ajout et l'identification de documents (photographies, cartes postales). Car avec les quatre premiers types de projets il y a une valeur ajoutée sur le document d'archives et parfois une production (documents, base de données, carnets, etc.). Pour les deux derniers types de projet, nous les avons choisis car ils contribuent à enrichir les fonds d'archives.

En plus les projets que nous avons choisis s'inscrivent dans la gratuité, le volontariat, la motivation et la bonne volonté des publics (internautes et lecteurs des salles). Aucun des projets ne fait l'objet d'une contrepartie financière entre les services d'archives et les publics.

Les différents types de projets choisis sont analysés, ci-après, dans les détails. Chaque type de projet sera étudié avec un exemple d'illustration. Ces exemples sont constitués par les projets des services d'archives de notre échantillon, sauf l'ajout de documents. Pour ce type de projet nous présenterons brièvement des exemples car les services d'archives de l'échantillon n'ont pas mis en place ce type de projet.

Les projets suivants ont été choisis à partir de la liste proposée par le Service Interministériel des Archives de France sur le Portail National des Archives, *France Archives*²⁹. Cette liste propose l'ensemble des services d'archives qui ont mis en place un ou des projets de *crowdsourcing*. Pour plus de précisions, nous avons eu à visiter les sites Web de chacun des services d'archives afin de vérifier l'existence de traces de ces projets mais aussi et surtout voir le fonctionnement et les modalités de participation pour le public. Pour certains projets, la participation requiert l'ouverture d'un compte, pour d'autres cela s'opère par un simple clic sur un lien vers la plateforme du projet. Cette authentification pour certains projets permet de récupérer des informations et de constituer une base de données. Sur le site des Archives de Bordeaux Métropole, la participation ne requiert ni compte ni clic, il s'agit de commenter tout simplement une proposition d'un autre participant, en remplissant un formulaire avec nom, prénom et commentaire sur la photo en question. Les modalités de participation sont donc très diverses d'un site à un autre et selon le type de projet.

1. Ajout de documents, de photographies, de cartes postales

Les pratiques participatives sur le Web peuvent être utilisées par les services d'archives pour l'accroissement des fonds. Cela consiste pour les internautes à ajouter des documents d'archives personnelles (lettres et photographies), ou des cartes postales. Ce type de projet participe en effet à enrichir les fonds d'archives du point de vue quantitatif. C'est le cas de la communauté GeneaNet (près de deux millions de membres), passionnée de généalogie, qui a mis sur son site Web le projet « Cartes postales ». Dans ce projet les internautes sont appelés à déposer et publier leurs propres cartes postales.

Les Archives départementales du Lot-et-Garonne ont aussi ouvert en 2011 une cartothèque départementale où les internautes peuvent déposer en ligne les fichiers de leurs cartes postales numérisées, mais aussi contribuer à l'indexation des cartes postales conservées par les Archives. De même, Europeana invite les internautes à numériser et télécharger sur le site lettres, cartes postales et photographies personnelles de la Première guerre mondiale, pour enrichir le projet *Europeana* 1914-1918 [Moirez, 2012, p. 190].

²⁹ Pratiques collaboratives : les internautes acteurs aux côtés des services d'archives. francearchives.fr [en ligne], 2018. Disponible sur : <<https://francearchives.fr/article/26288084>> (consulté le 05/04/2018)

2. Identification de documents iconographiques

En général c'est une tâche plus ou moins simple qui consiste à identifier des lieux, des personnes, des scènes, des sites, des monuments, des habitations, etc. à partir de la base d'images proposée sur le site Web du service d'archives. Il s'agit de décrire avec un ou des mots clés précis l'objet de l'image en question. Parfois c'est avec un formulaire de saisie s'il se trouve que l'image à identifier est plus ou moins complexe. Toutefois même si une image est déjà identifiée, les internautes peuvent toujours faire des propositions en contactant les administrateurs ou en commentant les propositions des autres contributeurs.

« Les techniques de *crowdsourcing* sont utilisées tout particulièrement pour l'identification de documents iconographiques, auxquels il est impossible d'accéder par un moteur de recherche s'ils ne disposent pas d'un minimum de données descriptives » [Moirez, 2012]. Plusieurs services d'archives français ont ainsi mis en place d'efficaces outils collaboratifs d'identification de photographies, soit sur leur site institutionnel comme aux Archives de la Haute-Garonne ou encore sous forme ludique aux Archives de l'Ain avec les enquêtes « SOS détective », soit sur des sites de partage comme Flickr, aux Archives des Alpes-Maritimes [Moirez, 2012, p. 189]. Outre une identification textuelle, le *crowdsourcing* peut permettre la géolocalisation de documents, comme l'interface ludique *Map Warper* de la *New York Public Library*³⁰ qui propose de superposer des cartes anciennes de New York à des cartes actuelles, et calcule ensuite automatiquement les données géographiques.

a. Archives départementales d'Indre-et-Loire³¹

Sur le site Web des Archives départementales d'Indre-et-Loire, l'identification consiste en une proposition. Pour participer il faut envoyer sa proposition par courrier, par téléphone ou par courriel via un formulaire. La personne proposant l'identification doit systématiquement mentionner la cote du document concerné.

Plusieurs thématiques sont proposées et enrichies au fil des acquisitions et du classement. Les documents proposés proviennent de la base *Collections de Touraine*, où leur notice complète est consultable de même que leur reproduction numérique de grand format. Les informations recueillies aideront ainsi à compléter la base de données iconographique du service.

³⁰ <http://maps.nypl.org/warper>

³¹ http://archives.cg37.fr/Chercher/RECONNAISSEZ_VOUS_CES_IMAGES_CONTRIBUEZ_A_LES_IDENTIFIER_!-ABCO.html

Les thématiques suivantes ont été relevées sur le site internet :

- Sites et monuments
- Coiffes tourangelles
- Scènes de rue
- Habitations
- Événements et rassemblements
- Individus



Figure 8 : page du projet d'identification d'images d'Indre-et-Loire

Ce service d'archives mène, dans le cadre du centenaire de la Grande Guerre, entre autres un projet d'édition de carnets de guerre et celui de transcription des carnets de guerre des poilus sur Wikisource.

b. Archives départementales de la Haute-Garonne³²

Ce service d'archives a proposé le projet d'identification du fonds « Charles Chevillot » sur son site Web institutionnel. Les inventaires de ce fonds sont systématiquement mis même en ligne au fur et à mesure du classement. Le fonds en question fait partie des « Fonds de personnalités locales ». Il rassemble les photos prises par Charles Chevillot (1891-1980) qui était un photographe passionné d'architecture qui a vécu à Toulouse. Ses photographies sont constituées d'éléments architecturaux de la ville : porche, cour intérieure, façade, etc. Le service d'archives a donc fait appel aux connaisseurs de la ville de Toulouse afin d'identifier ces photographies et de retrouver leurs adresses des sites en question. Les informations que donnent les participants laissent de précieuses indications qui permettant aux archivistes de légénder les images. La participation consiste à un commentaire sur l'image visualisée dans la fenêtre « identification de photographie ». Le contributeur est invité à saisir seulement son nom et son commentaire en bas des ceux des autres contributeurs.

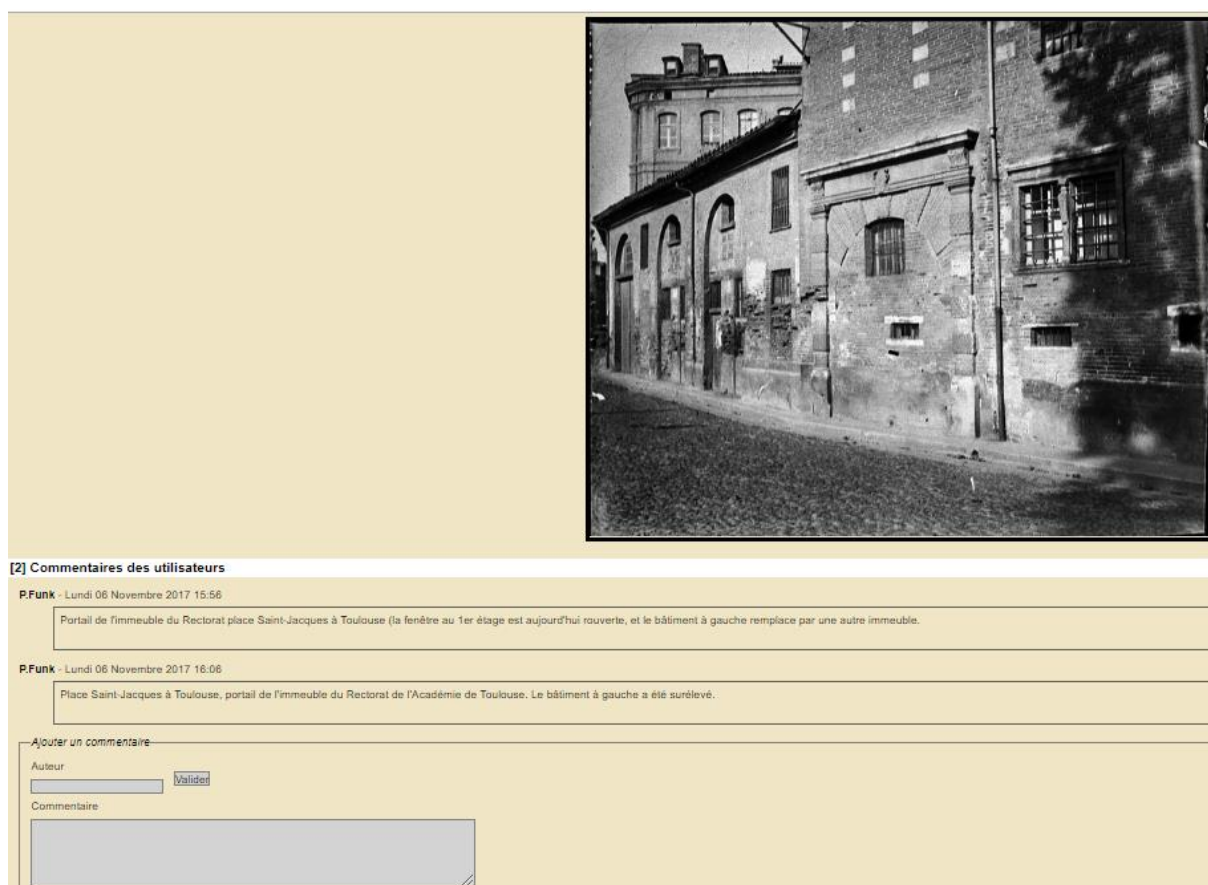


Figure 9 : page d'identification de photographie de la Haute-Garonne

³² http://www.archives.haute-garonne.fr/archives_en_ligne/archives_identifier.html

3. L'annotation collaborative

L'annotation est une opération qui consiste à la saisie de données, le plus souvent, dans un formulaire aux champs préétablis. Les champs du formulaire sont adaptés à chaque type de document. Ce type de projet peut parfois être similaire à l'indexation collaborative parfois en termes de procédure et de finalités. L'objectif étant de décrire les documents avec des données le plus souvent existantes, c'est-à-dire contenues dans les documents.

a. Archives municipales et communautaires d'Orléans³³

Sur *archives.orleans-metropole.fr*, le module d'annotation collaborative permet à l'internaute de saisir des données qu'il recopie à partir d'un document numérisé et sur la base d'un formulaire de saisie établi. Le but est de faire remonter des données en constituant une base de recherche instantanément interrogeable par mot par tout internaute.

L'annotation collaborative sur le site des archives municipales et communautaires d'Orléans consiste d'abord à créer un compte personnel si le volontaire n'en dispose pas encore. Ensuite l'internaute doit demander à participer à l'annotation. Après validation par l'administrateur du site, il peut alors se connecter à son espace personnel. Ainsi, une grille d'annotation apparaîtra quand les pages de certains fonds sont visualisées.

Les fonds ouverts à l'indexation sont :

- Registres paroissiaux constitués des registres de baptêmes, mariages et de sépultures des paroisses d'Orléans. Ce fonds comprend des tables communes aux différentes paroisses pour la période 1701-1792. Les tables ont été réalisées au XIXe siècle mais les actes les plus anciens remontent à l'année 1526 ;
- État civil comprend des tables décennales (1792-1922) et registres annuels des naissances (1793-1916), des mariages, des divorces et des décès allant de 1793 à 1926 ;
- Recensements de populations sont composés de listes nominatives pour la période de 1836-1911. Elles sont trouvables à minima via l'adresse Seuls les registres de plus de 100 ans sont en ligne et seuls ceux de plus de 120 ans peuvent être ouverts à l'annotation collaborative ;
- Actes d'engagement militaire sont des documents individuels d'engagement dans l'armée de la période 1818-1928. Ces documents ne concernent pas que des habitants

³³ <http://archives.orleans-metropole.fr/r/547/>

ou natifs d'Orléans. On y trouve aussi les actes d'engagement liés à la guerre de 1870, à celles de Chine et de celle de 1914-1918 ;

- Procès-verbaux et délibérations du Conseil municipal d'Orléans proviennent du Comité provisoire de la commune (1789-1790), du Conseil général de la commune (1790-1794), du Bureau et conseil municipal (1790-1795) et du Conseil municipal (après 1812).

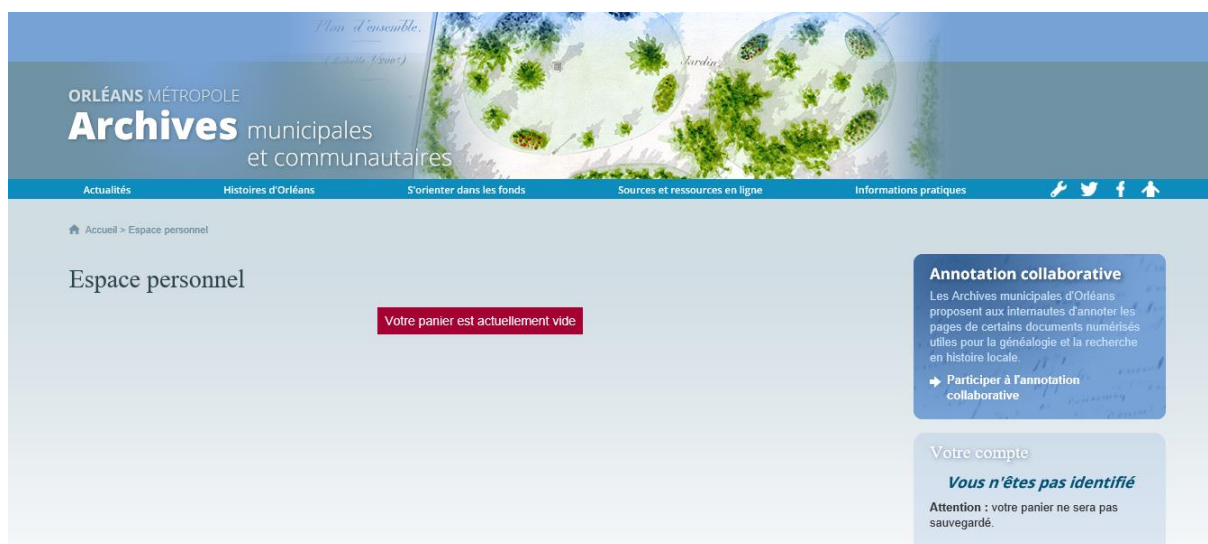


Figure 10 : espace personnel de l'annotation collaborative d'Orléans

b. Archives départementales de l'Aube³⁴

Ce service d'archives propose aux internautes un module d'annotation collaborative consistant à un référencement des patronymes présents dans les registres paroissiaux et l'état civil et les listes nominatives des recensements de la population. Les modalités de participation sont listées dans un document appelé "charte des annotateurs". Pour s'inscrire et participer comme annotateur au programme d'annotation, il faut disposer d'un compte personnel. Après cela il faut choisir selon ses centres d'intérêts un lot à annoter parmi ceux proposés par le service d'archives. L'accès au programme d'annotation collaborative est possible dans *MyArchive*³⁵, accessible depuis la page d'accueil du site internet. Ensuite pour demander l'ouverture de droits d'annotateurs, il suffit de cliquer sur le lien proposé dans son espace personnel. Au bout de quelques jours, la mention « Vous avez accès au module d'annotation » apparaît à droite de l'écran. L'annotateur peut, à titre exceptionnel, annuler un lot et le libérer pour que d'autres annotateurs puissent poursuivre le travail d'analyse. Lorsque l'annotation d'un lot est achevée,

³⁴ <http://www.archives-aube.fr/r/51/>

³⁵ http://www.archives-aube.fr/arkotheque/espace_personnel/

il faut le finaliser afin que les patronymes relevés soient mis à disposition de tous les internautes, via la base de données des patronymes.

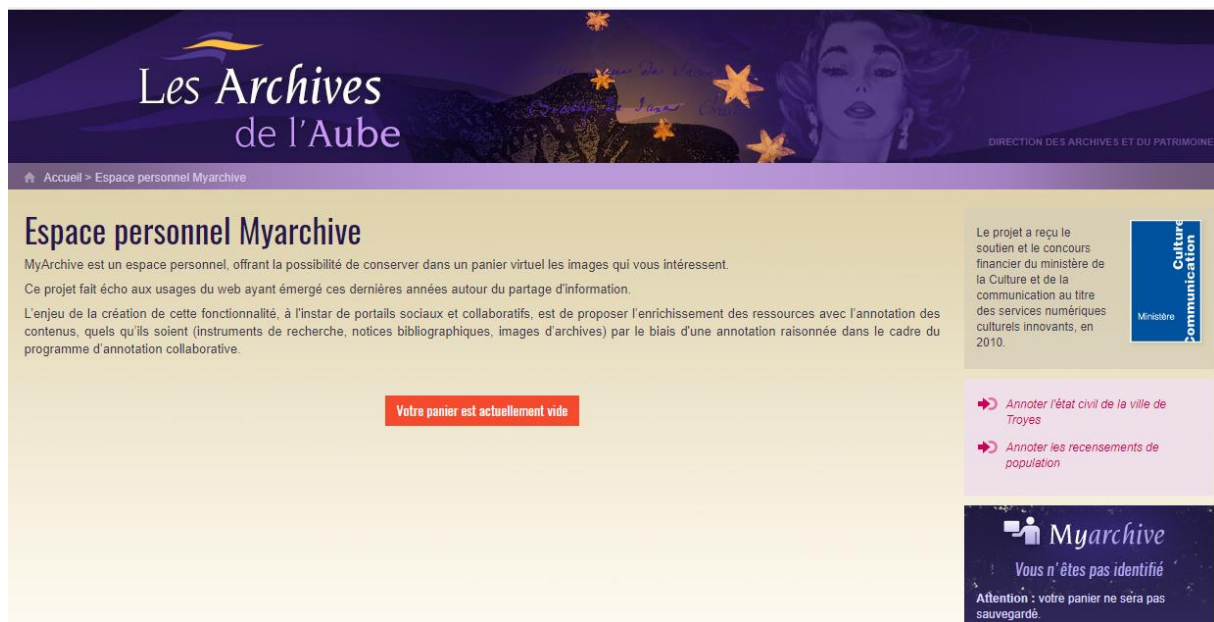


Figure 11 : espace personnel MyArchive pour l'annotation collaborative de l'Aube

4. Indexation collaborative

Le travail d'indexation peut parfois être assimilé à l'annotation. À vrai dire les deux activités nécessitent parfois la même procédure et adoptent les mêmes méthodes. L'indexation consiste précisément à l'association de mots clés à des documents d'archives identifiés, sélectionnées par les services. Ces documents peuvent être entre autres des images, actes numérisés, des cartes postales à titre d'exemple.

Cependant à la différence des *folksonomies* où l'utilisateur est laissé très libre de ses choix de mots clés, il est possible d'encadrer strictement les activités des usagers dans l'indexation. Ceci, pour permettre la constitution de bases de données structurées. C'est le choix fait par de nombreux services d'archives, en France pour le traitement de documents intéressant la généalogie par exemple. Une vingtaine de services d'archives français ont ainsi mis en place sur leurs sites web des modules d'indexation collaborative de documents nominatifs (état-civil, registres matricules militaires, recensements de population, etc.). Les chiffres sont parlants : aux

Archives départementales de l'Ain, 500 000 pages ont été indexées en 2 ans ; aux Archives départementales du Cantal, 1000 micro-tâches d'indexation sont réalisées chaque jour³⁶.

Ainsi, l'indexation reste parmi les activités les plus répandues dans les services d'archives en matière de projets collaboratifs.

a. Les Archives départementales de l'Ain³⁷

Les Archives départementales de l'Ain ont mis au point une plateforme d'indexation de l'état civil, des recensements de population et de listes électorales. Sur le site internet, ce sont des centaines de millions de pages numérisées prêtes à être indexées avec le volontariat des internautes. Ici l'indexation est un peu semblable à la transcription puisque la tâche demande un travail minutieux sur les manuscrits.

Pour faciliter le travail aux internautes, les Archives de l'Ain ont établi deux modes d'emploi : l'un pour l'indexation en général et l'autre pour le cas plus particulier de l'indexation des registres matricules. Dans ce cadre le processus pour l'internaute débute depuis la création de son espace personnel à la soumission de son indexation pour validation. Pour commencer à indexer l'internaute doit d'abord créer un espace personnel sur le site. Après la validation de l'inscription, il doit ensuite passer par l'onglet « archives numérisées » puis « état civil » et choisir au minimum une commune, de manière facultative, un type d'acte et une année et cliquer sur « Rechercher ». Enfin dans la liste des réponses, il peut choisir un lot d'images en cliquant sur « l'appareil photo » et ainsi ouvrir la visionneuse d'images. Cela lui permet de commencer l'indexation, en cliquant sur l'onglet « Indexation », puis sur « indexer cette image ».

³⁶ JOSSE, Isabelle ; MOIREZ, Pauline ; MOREUX, Jean Philippe. État de l'art en matière de crowdsourcing dans les bibliothèques numériques. Rapport L-4.3.1 du projet de R&D du Fond Unique Interministériel 12 [en ligne]. Paris : FUI, 2013, 77 p. Disponible sur : <http://www.bnf.fr/documents/crowdsourcing_rapport.pdf> (Consulté le 03/01/2018)

³⁷ <http://www.archives.ain.fr/n/pourquoi-participer/n:52>

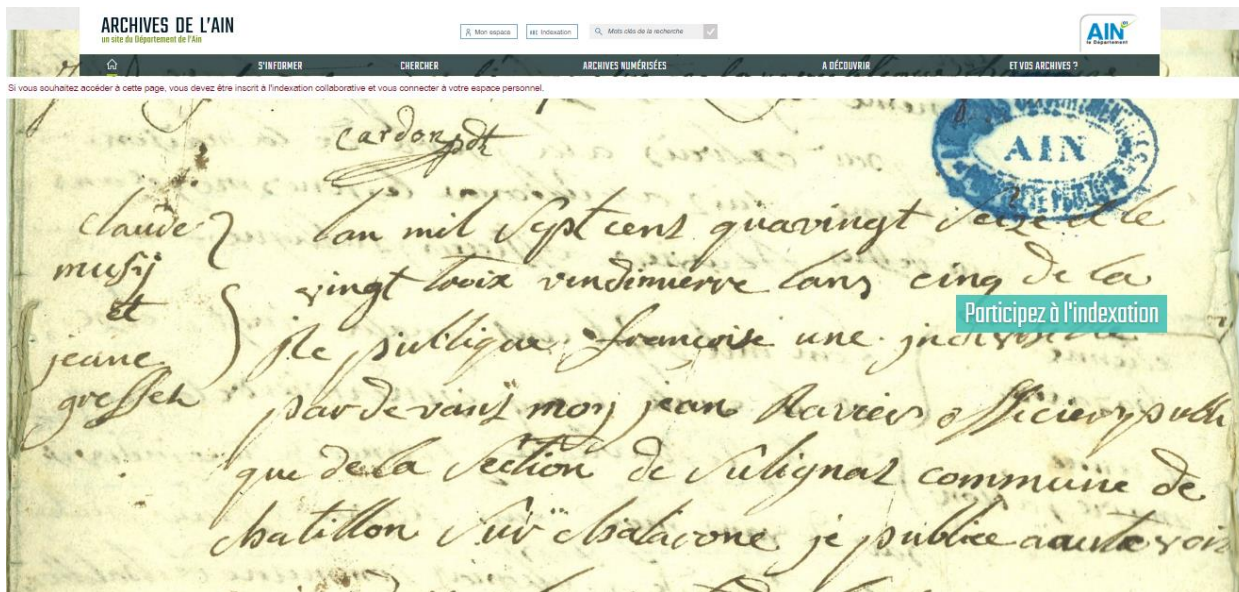


Figure 12 : page d'indexation des Archives de l'Ain

b. Ministère des Armées (Mémoire des Hommes)³⁸

Ce projet d'indexation collaborative est ouvert aux internautes afin d'annoter les fiches de la base des « Morts pour la France de la Première Guerre mondiale ». Cela permet d'enrichir les index déjà existants (nom, prénoms, date de naissance, département ou pays de naissance) et ainsi permettre des recherches plus fines parmi les 1,4 million de noms que comporte la base : recherche par lieu de naissance ou décès, par date de décès, par unité, grade ou encore classe et bureau de recrutement. Ce projet d'indexation collaborative des fonds est basé sur la gratuité, le volontariat, la motivation et l'esprit d'entraide des internautes. En devenant volontaire, l'internaute accepte implicitement les principes et s'engage à fournir le travail le plus rigoureux et le plus précis possible.

La participation à l'indexation collaborative des fonds nécessite la création d'un espace personnel. L'internaute doit par la suite demander à participer à l'indexation en acceptant les règles. Depuis le 9 novembre 2015, la validation de la demande de participation est immédiate. Une grille d'annotation apparaîtra quand les fiches des morts de la Première Guerre mondiale sont visualisées. Pour annoter à nouveau, l'internaute doit se connecter à son espace personnel à chaque nouvelle session. Il est aussi possible de suivre le nombre d'indexations réalisées à partir de l'espace personnel et connaître son rang d'annotateur. Cette dernière possibilité a été ajoutée depuis le 21 mars 2017.

³⁸ <http://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?larub=52&titre=annotation-collaborative>

The screenshot displays the 'Mémoire des Hommes' website interface. At the top, there is a navigation menu with categories like 'Présentation', 'Conflits et opérations', 'Présence française dans le monde', 'Parcours individuels', 'Matériels et équipements', and 'Musées et Collections'. Below the menu, a banner features the project title 'Mémoire des Hommes' and a call to action: 'Morts pour la France 14-18 : rendez hommage à un poilu et contribuez à la connaissance historique. Participez à l'indexation collaborative !'. The main content area is titled 'Indexation collaborative' and includes the following text:

Rejoignez le programme d'indexation collaborative et participez à l'enrichissement de la base des Morts pour la France de la Première Guerre mondiale

Le ministère des armées propose aux internautes d'annoter les fiches de la base des Morts pour la France de la Première Guerre mondiale afin d'enrichir les index déjà existants (nom, prénoms, date de naissance, département ou pays de naissance) et ainsi permettre des recherches plus fines parmi les 1,4 million de noms que comporte la base : recherche par lieu de naissance ou décès, par date de décès, par unité, grade ou encore classe et bureau de recrutement.

L'objectif est d'achever l'enrichissement de la base pour le centenaire de l'armistice 14-18, le 11 novembre 2018.

Ce projet d'indexation collaborative des fonds est basé sur la gratuité, le volontariat, la motivation et l'esprit d'entraide des internautes. En devenant volontaire, l'internaute en accepte implicitement les principes et s'engage à fournir le travail le plus rigoureux et le plus précis possible.

Si vous souhaitez participer à l'indexation collaborative des fonds :

1. Créez un espace personnel si vous n'en n'avez pas encore
2. Demandez à participer à l'indexation, en acceptant les règles
3. La validation à votre demande de participation est désormais immédiate (9 novembre 2015)
4. Une grille d'annotation apparaîtra quand vous visualiserez les fiches des morts de la Première Guerre mondiale. Vous pouvez annoter

On the right sidebar, it shows '1616108 annotations réalisées à ce jour.' and a 'Recherche globale' section with a search bar and the logo of the 'Direction des Patrimoines, de la Mémoire et des Archives'.

Figure 13 : projet d'indexation collaborative du Ministère des Armées

c. Archives nationales de France (Répertoires des Ponts et chaussées)³⁹

En janvier 2017, les Archives nationales ont lancé le projet collaboratif des Ponts et Chaussées consistant à relever, à partir des images numérisées présentées dans la salle des inventaires virtuelle, les informations contenues dans les tables et répertoires du Conseil général des Ponts et Chaussées constitués entre 1814 et 1965. Ainsi, les Archives nationales font appel aux internautes pour participer à l'enrichissement de ces inventaires. Ce projet permettra de créer une clé d'accès à l'ensemble des avis rendus par cette institution qui a en charge l'aménagement du territoire depuis le 18^e siècle. Dans la rubrique « Présentation du projet » nous avons pu relever en mai 2018 quatre articles qui nous informent sur le fonds d'archives concerné, le déroulement du projet, une foire aux questions ainsi que les missions ouvertes. Les projets participatifs des Archives nationales proposent d'indexer et de dépouiller des corpus de documents d'archives, à un niveau très fin, difficilement atteignable par les archivistes au regard du volume des fonds d'archives conservés. Les contributions des internautes sont intégrées aux inventaires et publiées dans la Salle des inventaires virtuelle. Interrogeables via le moteur de recherche, les contenus collaboratifs bénéficient à tous en permettant d'améliorer les résultats des recherches.

Le projet d'indexation collaborative des Ponts et Chaussées se décline en missions, elles-mêmes divisées en lots, permettant aux contributeurs de s'inscrire sur des lots précis à traiter.

³⁹ <http://archivnat.hypotheses.org/>

- Mission 1 : *De la chute du Premier Empire à la Restauration. 1814-1815.*
- Mission 2 : *Le Conseil général des ponts et chaussées aux changements de régime : 1870-1871.*
- Mission 3 : *Le Conseil général des ponts et chaussées pendant la Première Guerre mondiale. L'entrée en guerre, 1914.* Cette mission est achevée. (dernière mise à jour : 18/07/2017).
- Mission 4 : *Le Conseil général des ponts et chaussées pendant la Première Guerre mondiale. La guerre longue, 1915-1916.* Cette mission est achevée. (dernière mise à jour : 04/09/2017).
- Mission 5 : *La loi du 11 juin 1842 sur les chemins de fer. 1842-1843.*
- Mission 6 : *Le plan Freycinet de 1878. 1878-1882*
- Mission 7 : *Le Conseil général des ponts et chaussées pendant la Première Guerre mondiale. L'année de doute, 1917.* Cette mission est achevée. (dernière mise à jour : 23/10/2017).
- Mission 8 : « *Les premières années du règne de Louis XVIII, 1816-juin 1817.* »
- Mission 9 : « *1848 : De la Monarchie de Juillet à la Seconde République.* »
- Mission 10 : « *Le Conseil général des ponts et chaussées pendant la Première Guerre mondiale. La guerre en mouvement, 1918.* »

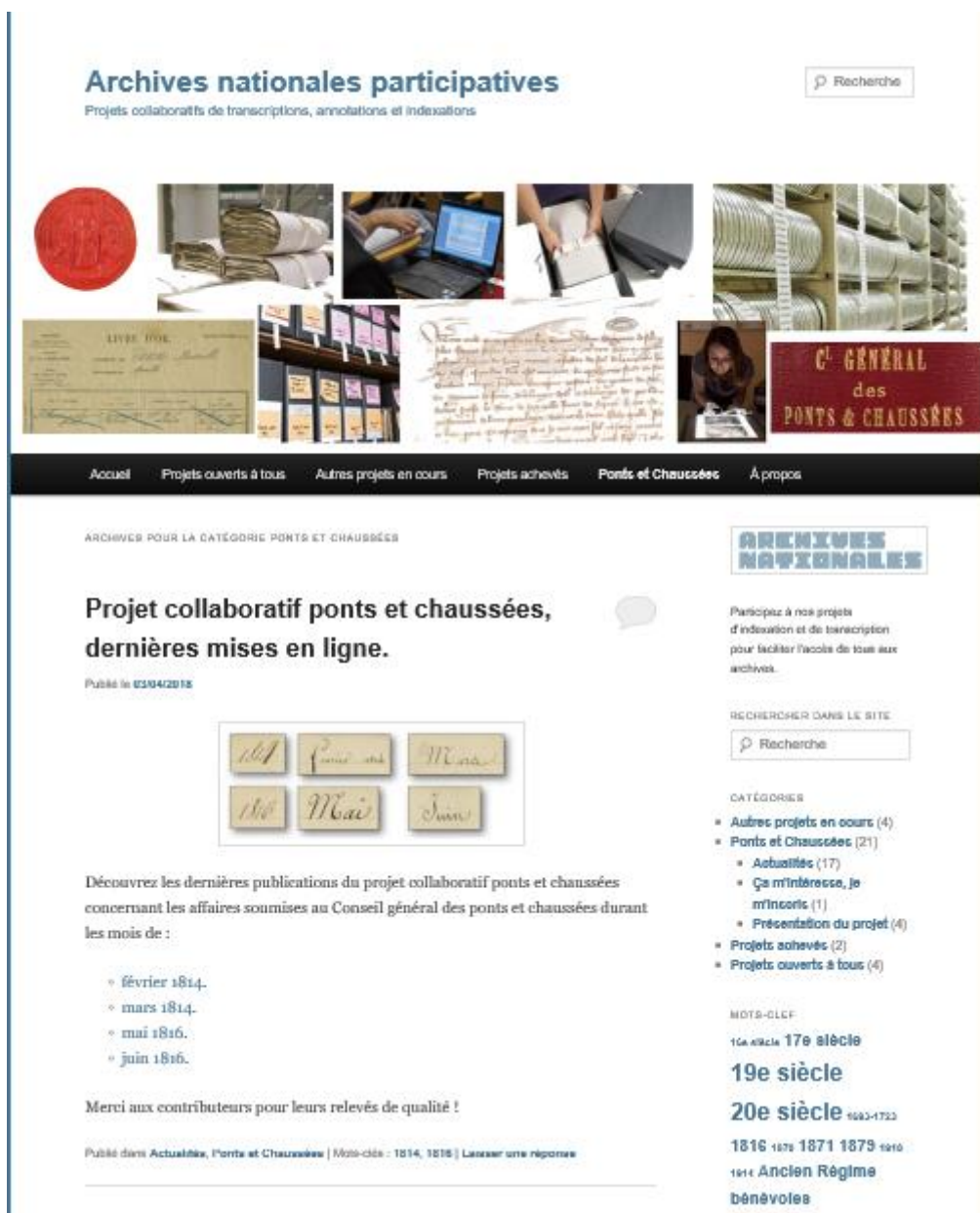


Figure 14 : page projet d'indexation collaborative des ponts et chaussées des Archives nationales de France

5. Transcription collaborative

C'est une tâche assez difficile qui demande beaucoup de sérieux de la part des internautes qui doivent relever chaque lettre et signe contenus dans les carnets ou testaments. Ces documents manuscrits, parfois pas entièrement lisibles, demandent beaucoup de concentration pour une transcription fidèle au texte d'origine. Les transcriptions sont par la suite éditées sous une forme numérique et mises en ligne.

La transcription collaborative peut s'effectuer sur des images numérisées seules ou accompagnées d'un OCR. Les internautes peuvent également être invités à corriger un texte préalablement OCRisé, ou même à le transcrire ex nihilo. Les techniques d'OCR automatique ne permettent pas en effet d'obtenir des résultats complètement parfaits, seule une relecture humaine permet d'atteindre un taux de reconnaissance de 100 %. De plus, l'OCR n'est à ce jour efficace ni sur les écritures manuscrites anciennes ni sur les livres imprimés avant le XVIIe siècle ; là encore, seul l'œil humain permet de réaliser une transcription de ces documents, afin de disposer d'un mode texte nécessaire à la recherche plein texte ou à la réalisation de livres numériques.

a. Archives nationales et Archives départementales des Yvelines (Testaments de Poilus)⁴⁰

Le projet « Testaments de Poilus » vise à produire une édition électronique d'un millier de testaments des Poilus de la Première Guerre mondiale retrouvés aux Archives nationales et aux Archives départementales. Pour cela, deux opérations sont nécessaires : transcrire le texte des testaments et l'encoder en XML TEI. XML TEI est un standard informatique utilisé notamment pour produire des éditions électroniques de documents manuscrits rigoureuses et faciles à exploiter. On ajoute au texte transcrit des informations sur la structure de ce texte, sur sa mise en page et sur la nature des objets qui le composent.

Les deux services d'archives ont ainsi fait appel à des volontaires du grand public pour réaliser ces tâches à bien. Pour cela ils ont élaboré un tutoriel expliquant les différentes étapes du travail. Les initiateurs ont aussi développé une plateforme participative sur laquelle les internautes peuvent transcrire les testaments et les encoder.

L'avancement de la transcription est signalé par un statut (à faire, en cours, en validation, validé)⁴¹.

⁴⁰ <https://testaments-de-poilus.huma-num.fr/#!/>

⁴¹ <https://testaments-de-poilus.huma-num.fr/#!/>

Figure 15 : page de recherche de testaments du projet de transcription collaborative

6. Edition collaborative de carnets de guerre

Cette pratique peut être complémentaire à la transcription collaborative. L'édition de journaux de guerre par exemple consiste à éditer sous une forme numérique les documents déjà transcrits. C'est donc grâce à la transcription collaborative que l'édition est possible. Certains projets d'édition sont effectués sur des plateformes wiki et ne demandent pas nécessairement d'inscription à l'internaute.

*Archives départementales du Pas-de-Calais*⁴²

Depuis le 2 août 2014, plusieurs services d'archives se sont associés dans un projet collégial d'édition de carnets de soldats de la Grande guerre sur Wikisource. Ces journaux de route, précieusement conservés par les descendants des poilus, ont été numérisés lors des deux grandes opérations de collecte menées en France : *Europeana* en 2013 et la *Grande collecte* en 2014. Les archives du Pas-de-Calais ont rejoint ce projet participatif et ont mis en ligne le premier carnet, celui d'Albert Labbé⁴³, jeune conscrit originaire de Pas-en-Artois, grièvement blessé à la tête en février 1915 en Champagne.

⁴² <http://www.archivespasdecalais.fr/Activites-culturelles/Chroniques-de-la-Grande-Guerre/Autour-des-commemorations/A-vos-plumes>

⁴³ https://fr.wikisource.org/wiki/Livre:Carnet_de_guerre_d%27Albert_Labb%C3%A9.pdf

Sur la page d'accueil du projet, le service donne l'accès à une page d'aide qui explique les différentes étapes pour participer et bien utiliser Wikisource.



Figure 16 : page d'accueil du projet d'édition des Archives du Pas-de-Calais

7. Catégorisation des projets

Puisque nous visons l'intelligence des publics nous avons choisi peu de projets d'ajout de documents. La plupart des projets des services d'archives répondants sont axés sur l'identification, l'annotation, l'indexation, la transcription et l'édition de carnets et de testaments de guerre. Nous allons catégoriser les projets selon 2 critères : le niveau de difficulté, l'évolution dans le temps. Cette catégorisation s'appuiera sur l'enquête, et sur les données trouvées dans la littérature et sur les sites Internet des services d'archives.

a) Répartition des projets

D'après les données de l'enquête, nous comptons 28 projets dans les 15 services d'archives répondants. Cela fait en moyenne d'à peu près deux projets (soit 1,86) par service d'archives. Certains services ont en effet plus d'un projet. Par exemple les archives départementales de l'Oise ont à elles seules 4 projets collaboratifs (indexation, annotation, identification, transcription).

Ce graphique ci-dessous fait une répartition des différents projets en fonction du nombre de fois qu'ils ont été réalisés dans les services d'archives. Ces derniers ne sont pas identifiés dans ce graphique car l'objectif est de voir les projets les plus réalisés dans ces services.

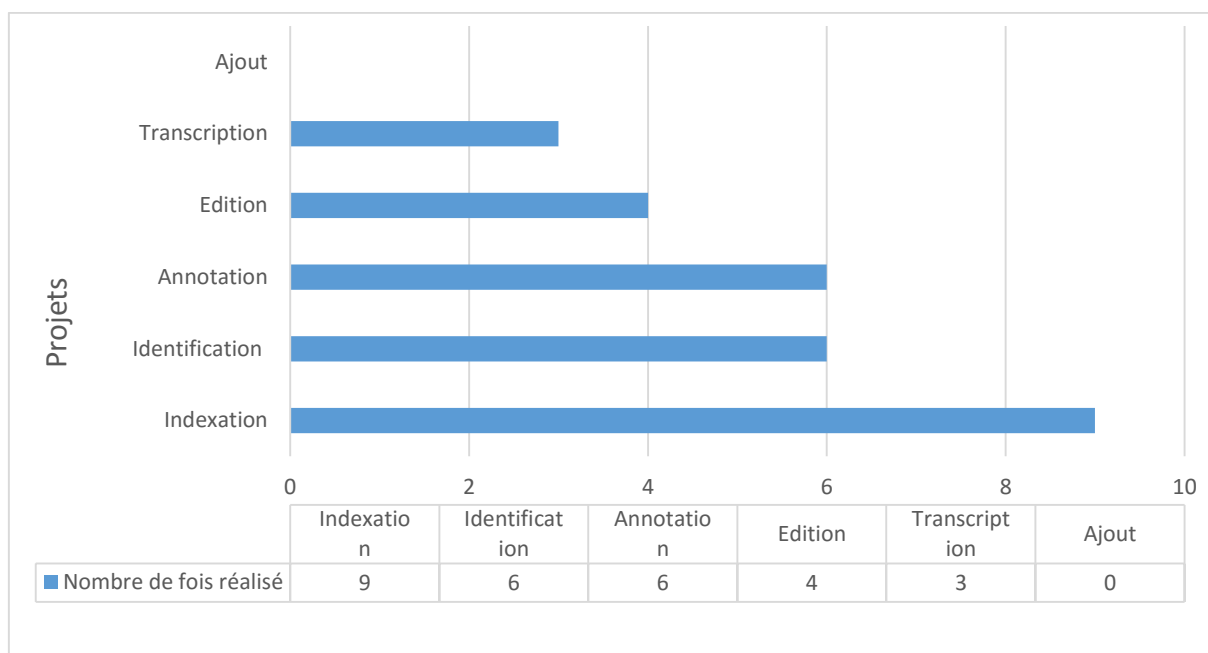


Figure 17 : Projets collaboratifs dans les services d’archives répondants

L’ajout de documents, comme nous l’avons remarqué, n’a pas été mis en place par les services d’archives. Néanmoins il fait partie de notre étude, c’est pour cela que nous l’avons représenté. Dans ce graphique nous constatons que dans les services d’archives, l’indexation collaborative domine largement avec 9 projets au total. En effet l’indexation collaborative a été un grand succès dans les services d’archives ces dernières années. C’est d’ailleurs la pratique du *crowdsourcing* la plus répandue dans les services d’archives. L’identification et l’annotation viennent en deuxième position avec 6 projets chacune. La transcription de texte n’est pas très représentée au niveau des services d’archives, seulement 3 projets ont été réalisés. L’édition de texte est aussi assez représentée dans les services d’archives avec 4 projets au total, malgré qu’elle est apparue tardivement (voir figure 20).

Le nombre important de projets d’indexation peut être expliqué par deux manières :

- soit c’est la forme la plus adaptée pour les fonds en question
- soit ces fonds en question sont plus nombreux dans les services d’archives.

Effectivement l’indexation portent toujours sur des documents en grand nombre dans les services d’archives : registres cadastrales, matricules militaires, recensement de populations, etc. Ces types de documents sont en nombre important à la différence des journaux de guerre qui sont en nombre réduit.

Le nombre est beaucoup plus élevé si nous l'associons à celui des projets d'annotation. Ces deux pratiques sont à peu près les mêmes, ce sont les appellations qui diffèrent d'un service d'archives à un autre. Notons quand même qu'il y a quelques différences notoires notamment la mise à disposition d'une liste de mots clés ou le choix de tags laissés aux internautes dans l'indexation, ce qui n'est pas le cas pour l'annotation.

Le petit nombre de projets de transcription peut être expliqué par la rareté de ces types de documents que sont les carnets et testaments de soldats. En effet ces documents sont précieux et jalousement conservés par les descendants des combattants. De ce fait ils ne sont pas aussi accessibles et nombreux comme les tables de l'état civil ou les registres paroissiaux qui sont des archives publiques. Nous pouvons aussi dire que le fait que le travail de transcription demande beaucoup plus de temps aux internautes peut expliquer ce nombre. De fait la nature du carnet fait que le temps de travail est beaucoup plus long que celui de l'indexation d'un registre d'état civil par exemple.

b) En fonction du niveau de difficultés

Selon le degré de difficulté du travail demandé à l'internaute, nous avons réparti ces projets en 5 niveaux ascendants :

- Le 1^{er} niveau concerne l'ajout de photographies, de cartes postales ;
- Le 2^{ème} c'est l'identification de photographies, de documents et/ou de lieux ;
- Le 3^{ème} niveau concerne l'indexation, l'annotation avec des mots-clés ;
- Le 4^{ème} niveau c'est la transcription de textes qui demande parfois des compétences particulières.
- Le 5^{ème} niveau c'est l'édition de carnets, de répertoires, etc., qui implique la création de documents, bases de données.

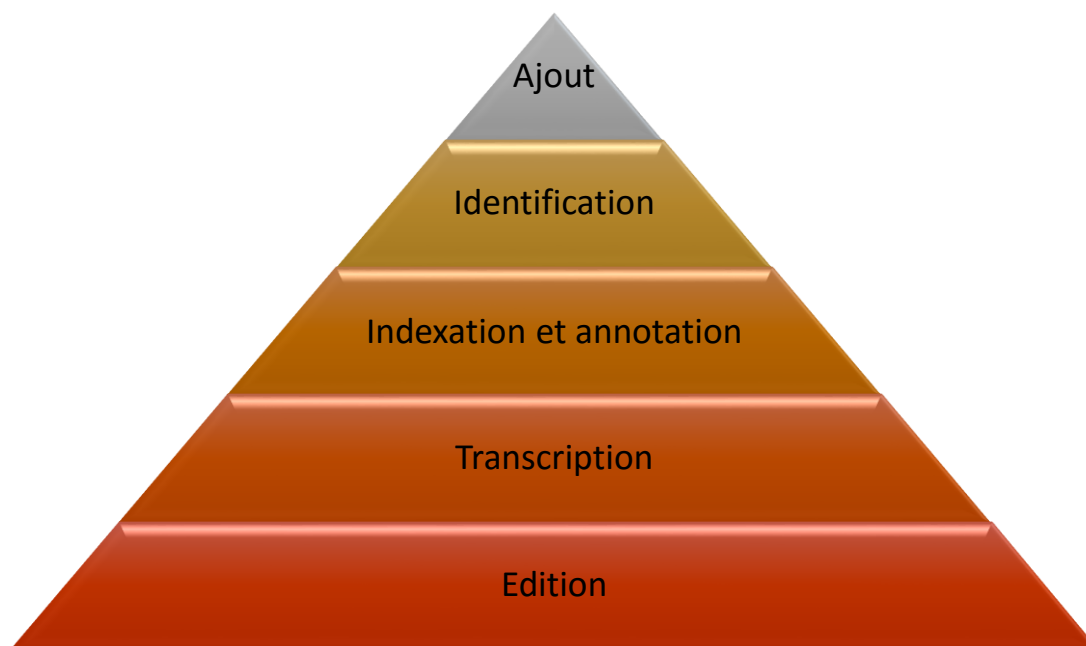


Figure 18 : Classement des projets de *crowdsourcing* selon le niveau de difficulté du travail de l'internaute

Dans cette représentation nous voyons nettement qu'à chaque niveau le travail demandé à l'internaute est de plus en plus sérieux et nécessite des connaissances et des compétences particulières. En effet selon ce classement, la transcription et l'édition deviennent, après l'indexation et l'annotation, les travaux les plus difficiles demandés aux internautes. Car elles demandent beaucoup de sérieux et de minutieux, de façon à relever lettre par lettre l'information contenue dans les écrits manuscrits. Ces deux types de pratiques demandent aussi beaucoup plus de temps aux internautes. Cela est lié à la nature des testaments et des journaux de guerre. De fait, ces documents sont premièrement rédigés à la main, deuxièmement assez long et troisièmement moins lisibles du fait du temps. C'est pour cela que le travail prend un peu plus de temps que dans d'autres opérations comme l'ajout de documents en guise d'exemple.

L'édition est encore plus difficile car elle implique la transcription d'abord et ensuite la mise en forme. C'est donc une opération à deux étapes qui demande beaucoup plus de moyens de la part des services d'archives : logiciels, mise en forme, type d'encodage, etc.

L'annotation et l'indexation viennent en troisième position dans ce classement car nous estimons qu'elles sont moins difficiles que les deux premières. Le travail demandé à l'internaute consiste à relever des informations parfois contenues dans les documents ou proposées dans une liste préétablie. En plus certains services d'archives proposent des formulaires de saisie, ce qui facilite davantage la tâche aux internautes. Par ailleurs le nombre important des documents

concernés par l’indexation et l’annotation peut demander beaucoup de participants. Cela implique une participation en masse dans ces opérations de *crowdsourcing*.

L’identification prend la quatrième position car à notre avis elle ne présente pas de difficultés notoires pour les internautes. Sa seule difficulté réside dans la connaissance de l’objet de la photographie à identifier. Le travail consiste uniquement à renseigner soit le lieu soit le bâtiment ou encore l’évènement qui se trouve sur la photo. Puisque l’identification porte souvent sur des photographies concernant une ville ou une localité, les connaisseurs peuvent facilement réaliser cette tâche.

Enfin l’ajout de documents vient en dernière position dans ce classement parce que c’est une opération très simple qui impose aux internautes d’avoir à leur disposition les types de documents demandés : cartes postales, photographies. La seule difficulté pourrait être le format et le type de résolution exigés.

c) En fonction de l’évolution des projets dans le temps

En se basant sur l’enquête, nous avons vu que depuis un peu plus d’une décennie (2007), les projets de *crowdsourcing* ont fait leur apparition dans les services d’archives. Dans le tableau suivant nous répertorions par ordre chronologique les différents projets suivant leurs dates de mis en place. Les données du tableau concernent seulement les réponses obtenues de l’enquête.

Année	Projets				Services d’archives
2007	Indexation	Annotation	Identification	Transcription	Archives départementales de l’Oise
2010	Annotation				Archives départementales du Cher
	Annotation				Archives départementales du Var
2012	Identification				Archives départementales de la Lozère
	Indexation	Identification	Transcription		Archives municipales et communautaires d’Orléans
2013	Annotation				Ministère des Armées
	Indexation	Annotation			Archives départementales des Yvelines

2014	Edition				Archives départementales du Pas-de-Calais
	Indexation				Archives départementales de l'Yonne
	Identification	Annotation	Indexation		Archives départementales de la Haute-Loire
	Indexation				Archives départementales du Cher
	Identification				Archives Bordeaux Métropole
2015	Edition	Transcription			Archives départementales d'Indre-et-Loire
	Identification	Edition			Archives municipales de Saint-Etienne
2016	Edition				Archives Nationales
	Indexation				Archives municipales de Saint-Etienne
2018	Indexation				Archives départementales du Pas-de-Calais
	Indexation				Archives départementales de l'Hérault

Figure 19 : Classement chronologique des projets des services d'archives répondants

Dans le graphique ci-dessous, suivons l'évolution des projets dans le temps, afin de percevoir pour chaque projet sa recrudescence et/ou sa décroissance. Cela nous permettra de connaître les projets les plus récents et les plus anciens, ainsi que ceux qui ont été le plus réalisés jusque-là. Ainsi, nous pourrons expliquer plus tard en fonction de cette analyse si en termes de valorisation les services d'archives utilisent les mêmes projets.

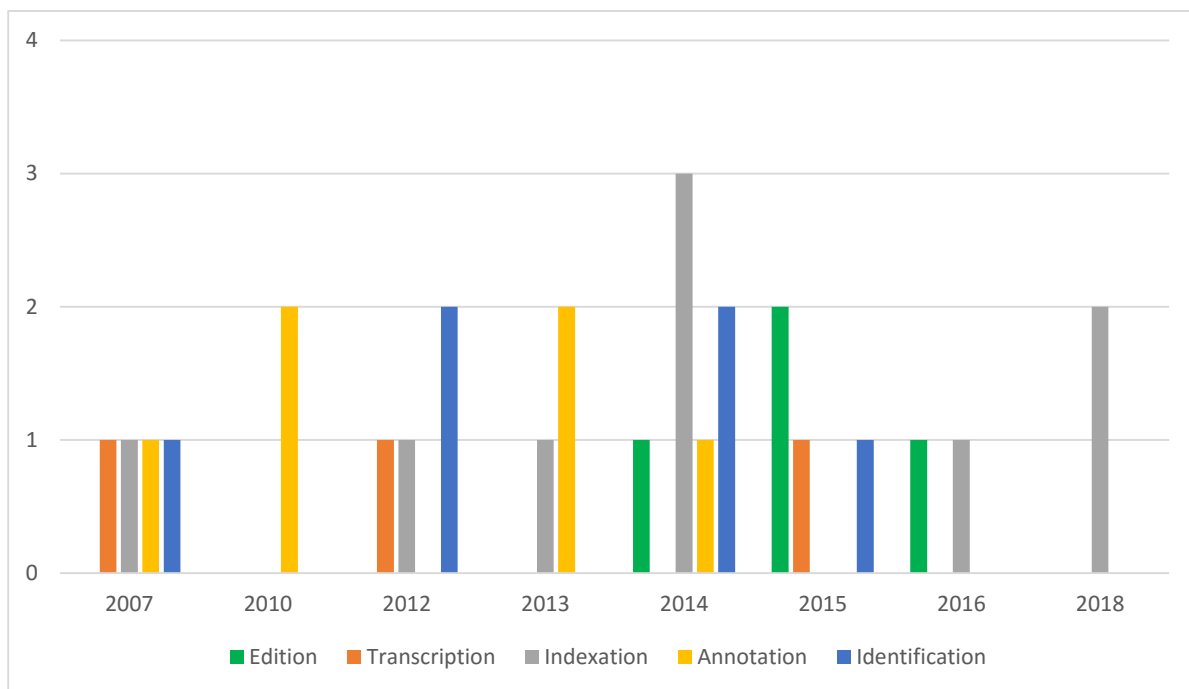


Figure 20 : Chronologie des projets mis en place par les services d'archives répondants

Dans ce graphique on voit que tous les projets sont représentés, sauf « l'ajout » de documents qui, comme précisé précédemment, n'est pas présent parmi les pratiques de notre échantillon de services d'archives. Nous constatons que seule « l'édition » est apparue plus tard, en 2014, aux Archives départementales du Pas-de-Calais. Les pratiques d'annotation dominent les autres avec 3 mises en place entre 2007 et 2010. Dans cette même période, pour chacun des projets de transcription, d'indexation et d'identification, il y a eu une seule mise en place. Il en restera ainsi jusqu'en 2012 où nous trouvons 2 projets d'identification. En 2013 les projets d'annotation l'emportent sur les autres. Ainsi, nous constatons que c'est toujours les mêmes pratiques depuis le début. Mais en 2014, il y a eu une innovation car nous trouvons pour la première fois un projet d'édition de carnets de guerre. À partir de ce moment aussi il y a eu beaucoup plus de mises en place de projets : 3 pour l'indexation et 2 pour l'identification. Les projets d'édition reviennent également en 2015 avec 2 réalisations. Entre 2015 et 2018, le nombre de projets n'évoluent pas beaucoup, 4 mises en place : transcription, identification, édition et indexation. Cette dernière reste au sommet en 2018, car c'est le seul projet que nous trouvons avec 2 réalisations. Si nous parlons en termes de nombre, nous constatons qu'il y a eu plus de projets d'indexation. La majeure partie des projets ont eu lieu entre 2014 et 2018. En effet c'est la période d'expansion des pratiques du *crowdsourcing* dans les services d'archives. Nous décomptons 15 projets dans cet intervalle contre 13 entre 2007 et 2014. Si nous parlons en

termes d'amélioration, nous remarquons qu'il y en a eu à partir de 2014. En plus des projets d'annotation, d'indexation, d'identification et de transcription, les services sont devenus plus innovants en proposant aussi l'édition de carnets de guerre. L'objectif étant d'instaurer ces pratiques collaboratives sur des types de documents rares et dans une forme plus élaborée. C'est donc une nouveauté car les manuscrits autrefois jalousement conservés par les familles pour la plupart, sont offerts au public. Dès lors ces derniers peuvent relever et saisir lettre par lettre ces manuscrits, afin d'en décortiquer le sens. C'est encore plus une innovation avec l'édition numérique et structurée avec le format d'encodage XML TEI.

III. LA VALORISATION DES ARCHIVES ET LE *CROWDSOURCING*

Dans un monde évolutif, marqué par des mutations technologiques ; et où l'information est créée toutes les secondes, il est impératif de rendre actuelles les informations passées avant qu'elles ne passent aux oubliettes. « On estime que tous les deux jours, l'humanité produit autant d'information numérique que pendant les deux millions d'années qui nous précèdent. » [Szoniecky, 2017, p. 11]. Dès lors, l'archiviste, gardien des traces de l'histoire, a la responsabilité de mettre en valeur les documents créés, qui parfois sont uniques. Cette mission de valorisation, comme toutes les autres, est affectée par les évolutions technologiques. En effet l'émergence du numérique a non seulement bouleversé les pratiques des archivistes mais aussi les habitudes des usagers. Aujourd'hui l'utilisation des outils numériques s'est généralisée à l'échelle mondiale et touche pratiquement tous les domaines d'activité. Le Web et les réseaux sociaux numériques font désormais partie du quotidien des utilisateurs. Il devient même difficile d'interagir avec ces derniers sans passer par l'Internet qui offre d'énormes possibilités en matière de diffusion et de dissémination de l'information.

Les projets collaboratifs sont le résultat d'une évolution technologique et de celle du Web, qui sont un moyen de promotion des fonds pour les archivistes. C'est alors une opportunité à saisir pour les archivistes qui désirent communiquer et diffuser leurs fonds. Cette large diffusion des documents d'archives sur le Web est en quelque sorte une forme de mise de valeur. En effet les archivistes se sont approprié ces facilités offertes par le Web social pour valoriser leurs fonds. Ainsi, les activités de *crowdsourcing* se sont beaucoup développées - d'une forme ou d'une autre - dans l'univers des archives, même si ses acteurs ont accusé un certain retard par rapport aux professionnels des bibliothèques et des musées. Ce décalage peut être compris dans la mesure où les stratégies et les objectifs visés diffèrent d'un type d'institution à un autre. Les services d'archives ont longtemps été dans un processus de numérisation de leurs fonds au moment où les autres institutions culturelles et patrimoniales multipliaient leurs actions de médiation.

Le *crowdsourcing*, comme nous l'avons précisé plus haut, a des impacts sur la valorisation des archives. Dans cette partie nous allons le démontrer en prenant l'exemple des projets identifiés dans notre enquête et en s'appuyant sur la littérature scientifique. Mais avant d'en venir à cette démonstration, définissons d'abord ce qu'est la « valorisation des archives ».

A. Qu'est-ce que la valorisation des archives ?

Cette mission est au cœur du métier d'archiviste qui doit l'assurer parmi tant d'autres qui lui sont assignées. Comprendre la signification exacte de cette fonction est donc indispensable pour l'archiviste quels que soit les changements de l'environnement. La valorisation c'est une activité essentielle pour les services d'archives. Elle pourrait être comprise comme une activité visant à rendre actuel des faits qui se sont déroulés dans le passé. Elodie Belkorchia⁴⁴ définit la valorisation comme étant « une actualisation des contenus qui peut aller de la simple éditorialisation au développement de projet pluri-média ». L'enjeu est « d'actualiser les traces anciennes pour qu'elles demeurent pertinentes en regard des préoccupations du présent sans dénaturer les significations originales » [Cardin, 2012, p. 40]⁴⁵. La notion d'actualisation est très importante à souligner. C'est le fait de rendre actuel, « faire passer de la puissance ou virtualité à l'acte ; c'est faire passer de l'état virtuel à l'état réel »⁴⁶. Dans ce même ordre d'idée, Jean-Luc Brackelaire⁴⁷, affirme que valoriser c'est virtualiser et valoriser contre la disparition. Par ailleurs il analyse la question d'un point de vue anthropologique en se fondant sur les études d'autres cliniciens et théoriciens. De ce fait, il soutient que la valorisation est un processus humain de base que chacun actionne pour pouvoir vivre la mutation incessante du présent et affronter la disparition et la perte. Adossant son propos à deux ordres de recherches psychologiques et anthropologiques, il définit :

« Valoriser c'est créer et faire entrer dans un nouvel ordre de réalité, celui du virtuel, où les choses s'actualisent en virtualité, en puissance, et adoptent le motif de l'intérêt, de la recherche, de la représentation, en lequel le présent à la fois se perd et re-présente profitablement. » [Brackelaire, 2012, pp. 21-31].

La valorisation consiste donc à une action de virtualisation de choses vouées à la disparition. Chez Jean-Luc Brackelaire la chose à valoriser est un objet de valeur et pour l'illustrer, il prend l'exemple du « doudou » du petit enfant. Pour lui c'est même le premier objet d'attachement des enfants. En cas de risque de perte, de perte ou encore d'abîme, l'enfant veut retrouver son « doudou » comme au début. En fait, ce constat vient d'un étonnement d'un couple de jeunes parents responsables d'une entreprise belge qui fabrique des peluches, des jouets et vêtements pour bébés et petits enfants. Les dirigeants de l'entreprise font part à l'auteur précité du fait que

⁴⁴ BELKORCHIA, Elodie, *opt. cit.*, p. 199

⁴⁵ CARDIN, Martine. In : HIRAUX, Françoise, MIRGUET, Françoise, *opt. cit.*, p. 40

⁴⁶ Centre National de Ressources Textuelles et Lexicales. Définition du verbe « actualiser ». Disponible sur : <http://www.cnrtl.fr/lexicographie/actualiser>. (Consulter le 12/05/2018).

⁴⁷ BRACKELAIRE, Jean-Luc. *Valoriser : virtualiser contre la disparition, une approche clinique et anthropologique*. In : HIRAUX, Françoise, MIRGUET, Françoise, *opt. cit.*

des parents leur contactent pour retrouver une peluche identique au « doudou » de leurs enfants. Ce fait n'est pas aussi banal qu'il paraît. Les parents comme les enfants sont dans le virtuel, c'est-à-dire qu'ils veulent retrouver à son état neuf un objet qui ne le sera jamais. Ils essaient alors de réaliser cette utopie en actualisant l'ancien objet. Pour y arriver ils n'hésitent pas à demander un coup de main pour retrouver un élément essentiel à la mise à jour de l'objet. Ainsi, ils font appel à des professionnels de la petite enfance pour rendre à l'état neuf le « doudou » qui vieillit ou pour retrouver l'identique de celui qui est perdu.

Retenons que la reconnaissance officielle est aussi très importante dans ce processus de mise en valeur dans la mesure où elle constitue une condition à toute forme de lien et de partage social. Sans l'appui des parents il serait très difficile pour l'enfant à retrouver son objet de valeur. Il ne pourrait pas identifier la peluche identique ni fabriquer un autre « doudou ». Cet état de fait avec l'exemple du « doudou » est aussi valable pour les archives. Valoriser les archives, nous dit Françoise Hiraux⁴⁸

« Passe aussi par la reconnaissance publique de leur importance dans le fonctionnement de la vie démocratique. Cette reconnaissance, à son tour, implique de poser une série de gestes concrets qui permettront d'organiser la préservation et la communication des archives en dépit de l'environnement qui leur oppose de grands obstacles. »

Par ailleurs elle revient sur la notion de valorisation des archives dans une approche communicationnelle. Pour elle, « valoriser des archives c'est transmettre et faire circuler des informations et des significations » [HIRAUX, 2012, p. 9]. Dans cette définition nous pouvons sous-entendre l'acte de médiation. Cela fait de l'archiviste un passeur d'informations. C'est également une mission de communication dans le fait de transmettre des informations. De plus nous pouvons dire que l'archiviste met tout en œuvre dans le but de faire parler les documents. Autrement dit, sans le travail de l'archiviste, les documents d'archives ne révèlent pas tous leurs secrets aux publics intéressés. Ce travail de valorisation est donc indispensable pour l'exploitation complète des archives.

Cela nous amène à nous intéresser aux liens qui existent entre la valorisation et les autres fonctions archivistiques.

⁴⁸ HIRAUX, Françoise, *Opt Cit.*, p. 15

1. La valorisation dans la chaîne archivistique

L'archiviste a certes d'autres fonctions à l'exception de la valorisation, mais celle-ci est une mission ubiquitaire. Elle se manifeste dans tous les maillons de la chaîne archivistique. C'est même la raison d'être de l'information. Une information non valorisée, peut être perdue, mal exploitée ou même inexploitable. Car comprenons le, toutes les actions archivistiques ont pour but de mettre en valeur les archives. C'est dans ce cadre que Martine Cardin⁴⁹ affirme :

« De loin en loin, toutes les actions d'un archiviste peuvent être considérées comme des actes de valorisation. C'est ainsi que la gestion des documents prévient la perte ou la destruction du patrimoine archivistique par l'encadrement des processus par lequel les documents acquièrent une valeur permanente. De même, l'acquisition conduit à statuer et à reconnaître cette même valeur. La classification et la description mettent en relief la portée et le contenu des fonds pour en favoriser l'exploitation. Quant à la communicabilité, elle structure les cadres d'accès intellectuel, physique et légal aux documents et fait la promotion des archives. »

Nous pouvons comprendre dans ces propos que la valorisation n'est pas une mission qui s'exécute une fois à un moment précis dans la chaîne archivistique. Elle peut s'effectuer implicitement au niveau des autres fonctions de l'archiviste. Elle est présente et pensée tout au long de la chaîne. Ces autres missions visent finalement la mise en valeur. C'est dans la perspective de mettre en valeur les d'archives que les archivistes s'appliquent dans l'acquisition, la description, le classement et la communication. C'est une mission centrale dans la chaîne archivistique. Nous pouvons alors comprendre qu'il existe une interdépendance entre les fonctions archivistiques.

Ainsi, Elodie Belkorchia⁵⁰ rejoint un peu les propos de Martine Cardin en affirmant que chaque maillon de la chaîne archivistique est interdépendant mais les liens ne semblent pas si évidents. En effet, la plupart des acteurs « (élus, public et parfois les archivistes) ne retiennent que l'élément qui est au bout de la chaîne, la valorisation. Ce qu'il faut aussi retenir c'est que cette valorisation peut éveiller la conscience du donateur (collecte) ou être une source de motivation pour les services versants de bien préparer leurs versements » Car, à travers la valorisation ils découvrent les efforts qui ont été faits sur le classement. Valoriser ses fonds c'est donc valoriser ses agents pour Elodie Belkorchia.

Ce qui est important à souligner c'est la relation qui existe entre la mission de valorisation et les autres fonctions archivistiques.

⁴⁹ CARDIN, Martine. In : HIRAUX, Françoise, MIRGUET, Françoise, *opt. cit.*, p. 33

⁵⁰ BELKORCHIA, Elodie : In la Gazette des Archives ; *Opt. Cit.*, p. 196

Par ailleurs la notion de mise en valeur exige des précisions aux niveaux théorique et pratique. Comprendre ce qu'est vraiment la mise en valeur des archives pourrait prévenir voire éviter de fausses guerres de territoire entre les experts du patrimoine écrit à savoir les archivistes et les historiens. En effet une querelle a eu lieu en 1980 entre des historiens d'entreprise et des archivistes d'une grande organisation financière québécoise⁵¹. Les premiers comme les seconds réclamèrent chacun de leur côté la garde des archives de l'entreprise. D'un côté les historiens croyaient être les premiers concernés par la mise en valeur des archives. De l'autre côté les archivistes ripostèrent en accentuant que les archives regorgent d'autres valeurs que les historiens n'étaient pas les seuls à utiliser. Effectivement, les archivistes devaient s'assurer qu'avec la diffusion, les archives atteindraient un plus large public. Fort de ce constat trois ordres de problèmes se révèlent selon Martine Cardin.

« Premièrement la polysémie de la notion de mise en valeur a été à l'origine de ce malentendu. L'expression « mise en valeur » est souvent employée pour désigner les notions d'exploitation, de diffusion, de référence et de valorisation. Deuxièmement certains considèrent que la valorisation des archives est un ensemble de moyens destinés à donner accès au corpus documentaire alors que d'autres estiment qu'il s'agit du fondement même de la mission archivistique. Troisièmement l'identification des valeurs en cause. La notion de mise en valeur a de nombreuses acceptions en archivistique : preuve, information, témoignage, valeurs primaire et secondaire, valeurs administratives, financières, légales, patrimoniales, auxquelles se juxtaposent certaines valeurs marchandes et esthétiques. Le chevauchement de toutes ces valeurs fut porteur de confusions, notamment sur la nature des activités de la mise en valeur et sur l'identification des clientèles à qui elle devait profiter. » [Cardin, 2012, p. 33]

Ce qui nous intéresse dans ce constat c'est la polysémie de la notion de « mise en valeur ». Elle peut recouvrir la diffusion, la référence et la valorisation. Toutes ces notions visent les mêmes objectifs selon Martine Cardin, même si les processus diffèrent et répondent chacun à un objectif spécifique. Le rôle de la valorisation c'est d'actualiser les traces du passé. La diffusion, elle, intervient sur la production récente, encore active. « La référence contribue à la sédimentation des strates anciennes et nouvelles d'archives en établissant des liens grâce à des instruments de contrôle intellectuel et administratif qui facilitent le repérage ». La référence participe donc à consolider les liens entre les informations passées à valoriser et celles récentes à diffuser. Mais qu'en est-il de ces deux fonctions ? La valorisation et la diffusion sont-elles divergentes ou convergentes en matière d'objectifs et d'actions ? Dès lors il est important de poser les similitudes et différences entre la valorisation et la diffusion, afin de comprendre leurs

⁵¹ CARDIN, Martine, *Opt. Cit.*, p. 33

liens et les leurs rôles dans la chaîne archivistique. Ce qui nous mène à nous poser la question suivante.

2. Diffuser et valoriser : quelles différences ?

Les auteurs qui ont voulu faire l'état de question sur la valorisation des archives sont souvent confrontés à des problèmes de terminologie. « Le vocabulaire est différent de part et d'autre de l'Atlantique » [Cardin, 2012, p. 34]. Les acceptions de la valorisation ne sont pas unanimes au niveau mondial. D'un côté en Europe c'est le terme de valorisation qui est utilisé, et de l'autre côté les québécois parlent plutôt de diffusion. Même si les termes diffèrent, les significations et les objectifs visés restent les mêmes.

Le terme de valorisation est utilisé en Europe pour désigner des actions par lesquelles les archivistes font connaître les archives et en facilitent l'exploitation. Ce terme n'est pas défini par le *Dictionnaire terminologique* des Archives de France⁵². C'est plutôt l'expression « mise en valeur et exploitation » qu'on retrouve dans ce dictionnaire. C'est dans un tableau méthodique qui esquisse une ontologie que ce concept est associé à l'animation culturelle, l'activité éducative, la communicabilité et d'autres modalités de valorisation telles l'exposition et la publication [Cardin, 2012 p. 35]. Cette acception correspond à celle du PIAF⁵³ qui consacre un module de formation entier à la valorisation des archives. Dans ce module il s'agit d'actions visant « l'exploitation pédagogique et culturelle des archives dans le prolongement naturel de la communication des documents »⁵⁴. En tant que système de pratiques, cette valorisation rejoint trois enjeux sociaux, à savoir : éduquer, témoigner et cultiver et elle se réalise à partir de quatre domaines soit : l'action éducative, l'exposition, les publications et la communication des archives.

Pour leur part, les Québécois n'utilisent pas le terme de valorisation mais plutôt celui de « diffusion » qu'ils définissent comme étant :

« L'action de faire connaître, de mettre en valeur, de transmettre ou de rendre accessibles une ou des informations contenues dans des documents d'archives à des utilisateurs (personnes ou organisations) connus ou potentiels pour répondre à leurs besoins spécifiques »⁵⁵.

⁵² Dictionnaire terminologique archivistique : Disponible sur : https://francearchives.fr/file/4575c619ab1e1e738d81d2249ff8dd4115a3d8cb/ARCHIVES_DE_FRANCE_Dictionnaire_de_terminologie_archivistique.pdf (consulté le 02/04/2018)

⁵³ Portail International Archivistique Francophone

⁵⁴ PIPON, Brigitte ; LAUBIE, Xavier. Module 12-Valorisation des archives. PIAF [en ligne]. Disponible sur : <www.piaf-archives.org/se-former/module-12-valorisation-des-archives> (consulté le 01/04/2018)

⁵⁵ CHARBONNEAU, Normand. La diffusion In Couture, Carol (dir.) et coll. *Les fonctions de l'archivistique contemporaine*. Québec : Presses de l'université du Québec, 1999, p. 374

Cette définition tirée d'un manuel sur les fonctions archivistiques résulte d'une volonté de normalisation de la terminologie québécoise, contrairement à la perspective européenne qui a davantage recours à une appellation courante dans la communauté. [Cardin, 2012]

Indépendamment de ces appellations, la valorisation et la diffusion semblent recouvrir les mêmes activités dans les deux traditions. Par ailleurs il faut reconnaître que diffuser et valoriser ne sont pas des synonymes. La muséologie et les études patrimoniales permettent de faire quelques explications. Pour celles-ci :

« La valorisation n'est pas un ensemble de moyens et techniques pour présenter de belles pièces choisies, mais une intervention véhiculant des propositions pour donner à penser le monde. La valorisation du patrimoine a pour but de produire un gain de connaissances et une transformation d'un sujet. Elle maintient vivantes des significations associées à des biens culturels menacés de destruction ou en perte de pertinence » [Cardin, 2012 p. 36].

Cette définition nous rappelle celle de Jean-Luc Brackelaire, qui précisait que valoriser c'est virtualiser et valoriser contre la disparition. « Valoriser suppose donc un engagement et une participation active des agents culturels dans un processus de construction de significations qui transcendent l'objet à valoriser pour actualiser les représentations qu'il médiatise » [Cardin, 2012 p. 36]. Cette acception et l'argumentaire des auteurs du module du PIAF se recourent. Brigitte Pipon et Xavier Laubie⁵⁶, auteurs du modèle, affirment :

« Valoriser les archives, c'est agir concrètement pour faire changer leur statut. De documents conservés et individuellement communiqués, voici qu'ils deviennent des documents mis en avant, choisis pour être rendus attractifs.»

La question qui nous interpelle après cette analyse est : laquelle de ces deux fonctions joue le *crowdsourcing*, surtout dans cette ère du numérique ? Autrement dit est-ce utilisé comme moyen de valorisation ou comme moyen de diffusion des archives ? La réponse qui semble correcte est que le *crowdsourcing* sert à valoriser les archives, alors que c'est plutôt la numérisation qui assure la fonction de diffusion des archives sur le Web. Ici nous entendons par « diffusion des archives » la dissémination massive sur internet. En outre la numérisation conditionne le *crowdsourcing* car s'il n'y avait pas cette mise à disposition des archives sur le Web, il n'y aurait pas de pratiques collaboratives autour des archives. C'est donc une condition indispensable pour instaurer des pratiques collaboratives autour des fonds avec les publics.

⁵⁶ PIPON, Brigitte ; LAUBIE, Xavier Module 12-Valorisation des archives. PIAF [en ligne]. Disponible sur : <www.piaf-archives.org/se-former/module-12-valorisation-des-archives> (consulté le 01/04/2018)

3. Forces des services d'archives : masse documentaire, documents uniques et aux valeurs multiples, etc.

La numérisation et la mise en ligne des archives historiques constituent l'une des méthodes de diffusion mais aussi de valorisation les plus répandues à l'ère du numérique. Cette diffusion des documents d'archives sur le web s'est largement démocratisé et a ainsi atteint un public considérable. Outre cette forme de communication, les archivistes ont utilisé les moyens offerts par le Web 2.0 pour mettre en valeur leurs fonds en profitant du nombre important d'internautes et de la pertinence des documents d'archives.

En effet les documents d'archives sont appropriés aux pratiques collaboratives en raison de plusieurs caractéristiques par lesquels ils donnent aux services d'archives d'énormes opportunités. Ainsi, Pauline Moirez affirme : « le volume des fonds d'archives fournit aux services d'archives une matière première abondante et adaptée à la diffusion numérique. »

La masse documentaire est donc l'une des forces que détiennent les services d'archives. À la différence des autres institutions culturelles et patrimoniales, les services d'archives disposent d'un volume de documents important qui s'accroît sans cesse. Cela est dû au fait que les institutions, les administrations, les entreprises (publiques et privées), les banques, bref tous les organismes produisent chaque seconde des milliers de documents dont une partie deviendra des archives. En plus aujourd'hui on voit l'émergence d'outils numériques qui créent des volumes de documents inestimables avec l'ère du *big data*, sans parler des appareils mobiles. C'est donc une spécificité des archives qu'elles ne partagent même pas avec les bibliothèques, malgré l'émergence de la publication numérique et les maisons d'édition qui se multiplient. Cette masse de documents comme l'indique Pauline Moirez, est adaptée à la diffusion sur le Web. En effet ces pièces d'archives sont adaptées à la numérisation et à la diffusion sur Internet. C'est une véritable matière première qui permet de concevoir des projets collaboratifs afin de disséminer les documents et d'interagir avec les publics.

Cette problématique de la masse peut expliquer parfois le décalage avec les réalisations de médiation numérique et de présence sur le web social d'autres domaines culturels. Jusqu'à une date récente la priorité des archivistes restait la numérisation et la mise en ligne des documents, au moment où les musées par exemple vont centrer leurs efforts sur la médiation et l'interaction avec leurs usagers quand bien même ils fournissent très peu de contenus en ligne. [Moirez, 2012, p. 188]

Cette force des services d'archives selon Pauline Moirez réside par ailleurs dans la nature des documents d'archives uniques, leurs valeurs multiples, les stratégies de mise en valeur. Dans ce cadre, elle affirme : « L'analyse « organique » du document d'archives explique le fait que des efforts contributifs peuvent être entrepris avec succès dans le but de mettre en valeur les fonds d'archives. » [Moirez, 2012, p. 188]. Le document d'archive est parmi les plus rares objets informationnels qui détiennent cette valeur d'unicité. À la différence d'un fonds documentaire, les documents d'un fonds d'archive sont inséparables s'ils veulent garder leur signification contextuelle. L'archive n'a de sens que dans son contexte de production associée aux pièces avec lesquelles elle a été produite. Cela fait que les projets contributifs dans les archives sont toujours porteurs de succès s'ils visent au final la valorisation. Car ils portent sur un ensemble cohérent, constitutif de documents.

Autre atout que portent les archives, leurs valeurs multiples. En effet les archives ont beaucoup de valeurs qui favorisent des usages multiples et variés. Dans cette optique Pauline Moirez souligne : « le contenu et la nature discursive des documents d'archives expriment des valeurs, historique, citoyenne ou encore éducative, qui génèrent des usages innovants, participatifs ou interactifs » [Moirez, 2012, p. 190]. Par conséquent la valorisation des archives peut également recouvrir plusieurs significations. De ce fait Martine Cardin atteste :

« La notion de mise en valeur a de nombreuses acceptions en archivistique : preuve, information, témoignage, valeurs primaire et secondaire, valeurs administratives, financières, légales, patrimoniales, auxquelles se juxtaposent certaines valeurs marchandes et esthétiques. »

Le chevauchement de toutes ces valeurs peut parfois être porteur de confusions, selon Martine Cardin, notamment sur la nature des activités de la mise en valeur et sur l'identification des clientèles à qui elle devait profiter [Cardin, 2012, p. 33]. Contrairement à cette auteure qui considère ces valeurs comme source de confusions, notamment dans la valorisation, Pauline Moirez en voit plutôt un moyen de multiplier les activités de mise en valeur. Cela permet en effet aux archivistes de disposer d'un large choix d'arguments pour initier des pratiques collaboratifs. Cela peut aussi expliquer la diversité des publics qui participent aux projets ayant des motivations très variées. Celles-ci peuvent aller de la simple besoin d'information à un objectif scientifique en passant par un désir de service à la communauté.

Les archivistes ont donc profité de toutes ces opportunités offertes par les documents d'archives pour mettre en place de nouvelles stratégies de mise en valeur de leurs fonds en développant

des projets participatifs. Ces projets participent donc à valorisation des archives par diverses formes.

B. La valorisation des archives par le *crowdsourcing* : comment ?

Afin de montrer comment le *crowdsourcing* participe-il à la valorisation des archives, nous allons voir la manière dont s'opère cette valorisation, comment elle se manifeste et quelles sont les retombées des projets sur les fonds d'archives. Ainsi, nous allons d'abord analyser l'évolution des projets, pour voir si elle a des impacts sur la valorisation. Dans cette partie nous allons essayer de démontrer les projets qui semblent le mieux valoriser les fonds. Ensuite nous allons enchaîner avec les types de publics : en se posant la question à savoir si est-ce les mêmes publics qui participent dans ces projets ? Les projets de *crowdsourcing*, ayant eu un grand succès ces dernières années au niveau des services d'archives, nous allons ainsi nous intéresser à leurs moteurs de réussite afin d'en identifier aussi les leviers de la réalisation de la valorisation. Nous terminerons enfin sur les finalités des projets de *crowdsourcing*, leurs retombées sur les fonds d'archives.

Les projets de *crowdsourcing* participent beaucoup à la valorisation des archives à travers diverses formes. Certaines formes sont développées à partir de l'exploitation des archives et sont :

« [...] de nouvelles écritures sur les archives, usages et pratiques avec à l'appui des QR-codes, d'applications mobiles, etc. »⁵⁷. Toutes ces nouvelles pratiques basées sur l'exploitation des archives contribuent à mieux les valoriser auprès des publics. Ce sont donc des moyens de mise en valeur efficaces que l'archiviste créatif met en place au moment où les lecteurs sont présents sur Internet et se font rares dans les salles de lecture. »

À la différence des bibliothèques et des musées, les projets participatifs dans les services d'archives sont essentiellement centrés sur le document d'archives. Les stratégies des services d'archives et celles des autres institutions culturelles et patrimoniales sont différentes, de même que les objectifs. Car comme le constate Pauline Moirez :

« Contrairement aux musées, qui peuvent utiliser le web pour accroître la fréquentation de leurs institutions, les services d'archives ne recourent pas au web social pour amener plus d'usagers dans les salles de lecture physiques. Pour eux il s'agit plutôt d'initier un travail scientifique collaboratif, à distance, sur la base d'une dissémination massive de leurs fonds d'archives » [Moirez, 2012, p. 192].

⁵⁷ Christelle BRUANT in Les mutations du métier d'archiviste et de son environnement, 2015, p.

Il en découle de ces travaux collaboratifs axés sur les fonds d'archives, une grosse production de données de qualité. Cela est possible grâce à l'encadrement qu'assurent les archivistes. Dans ce cadre Pauline Moirez affirme :

« Avant tout intéressés par la mise en valeur des documents d'archives et par l'enrichissement de leurs descriptions, les archivistes sont tout particulièrement sensibles à la qualité des informations produites par les usagers. Cette crainte reste d'ailleurs souvent l'un des principaux freins à la mise en place de ce type de projets. C'est pourquoi les opérations de crowdsourcing sont encadrées, et le plus souvent intégrées sur les sites web institutionnels plutôt que déportées sur les médias sociaux où les vérifications sont plus complexes à effectuer. »
[Moirez, 2012, p. 193]

Cela est en effet une nouvelle forme d'exploitation des documents d'archives, qui, à la différence des formes classiques, participe beaucoup à valoriser les fonds. Contrairement à la consultation sur place dans les salles de lecture, les pratiques du *crowdsourcing* permettent aux internautes, par exemple en transcrivant un testament de soldat, de l'exploiter en même temps. De même lors de l'indexation d'un registre de données nominatives (recensements de population, état civil, etc.), l'internaute peut avoir des renseignements sur les patronymes qui l'intéressent. Ainsi, les internautes qui participent à ces projets peuvent simultanément exploiter les documents d'archives.

Cette valorisation peut s'opérer selon deux modalités :

- Les archivistes, en mettant en place les projets de *crowdsourcing*, visent souvent la valorisation d'un fonds spécifique, souvent non connu du public ou partiellement. En proposant certains fonds, ils visent à les mettre en valeur ; par exemple le carnet d'un poilu numérisé (Archives départementales d'Indre-et-Loire)
- Les publics (internautes), en contribuant à décrire les fonds, participent aussi à leur valorisation auprès d'autres publics car ils transcrivent, ajoutent des mots clés, identifient et indexent les documents dans le but de leur donner plus de valeur. Cela participe ainsi à faciliter la recherche et à atteindre un public plus vaste que celui initialement visé.

Cette valorisation est facilitée par 3 éléments :

- Les supports des projets (sites web institutionnels, réseaux sociaux, plateformes dédiées aux pratiques du *crowdsourcing*) qui peuvent rassembler ou non un certain nombre de participants ;

- Les publics qui sont motivés à des degrés différents en fonction de leurs intérêts pour les fonds ;
 - Les fonds proposés dans ces projets collaboratifs.
1. Évolution des pratiques du *crowdsourcing* : quels impacts sur la mise en valeur des archives ?

Les pratiques du *crowdsourcing* sont de plus en plus présentes au niveau des services d'archives. Ces pratiques ont aussi évolué dans le temps, bien vrai que leur expansion ne datent que d'un peu plus de dix ans dans les services d'archives. Les raisons pour lesquelles les services d'archives sont en retard par rapport aux musées et aux bibliothèques, ont été évoquées dans différentes parties de ce mémoire.

Les premiers projets d'envergure au niveau mondial étaient centrés sur la numérisation et l'ajout de documents photographiques notamment aux Pays-Bas⁵⁸, en Allemagne⁵⁹, aux États-Unis⁶⁰ et en Angleterre. En France, plus précisément dans les services d'archives, ce fut en 2007⁶¹ que les Archives départementales de l'Oise proposent quatre types de pratiques collaboratives sur ses fonds. Ces pratiques étaient centrées sur la description des fonds : identification indexation, annotation, transcription.

La valorisation des fonds par ces projets dépend de plusieurs facteurs. Certes cette évolution a joué un rôle important avec l'édition, mais nous verrons que ça dépend aussi du projet et du type de fonds que les services d'archives désirent mettre en valeur. Nous constatons qu'il y a une relation de cause à effet entre les projets et les fonds en ce qui concerne la valorisation. C'est-à-dire que l'impact (valorisation, visibilité, etc.) des projets sur les fonds dépend de l'adaptabilité entre ces deux éléments.

Les premiers projets, apparus depuis 2007 dans certains services d'archives et ont été remis en place plus tard par d'autres. Mais nous avons vu que, de loin des autres projets, l'édition de texte est apparue un peu plus tard, en 2014 dans les services d'archives (voir figure 20). Même si elle est assimilable à la transcription, notons qu'il y a bien une différence entre ces deux

⁵⁸ Amsterdam, *You ask we scan*, 2009

⁵⁹ Wikimedia et Bundesarchiv, *Wikimedia Commons*, 2008

⁶⁰ National Archives, *Yes We Scan*, 2011

⁶¹ Cf. Figure 8

pratiques. En plus d'être de la transcription de texte manuscrit, l'édition structurée en TEI⁶², vise à produire une édition électronique de milliers de documents. C'est l'encodage en XML qui procure à la transcription un côté innovant. Ainsi, c'est une nouvelle forme de médiation numérique qu'explorent les services d'archives, à l'instar des Archives Nationales. Dans ce cadre : le service déclare que le but de son projet d'édition est de :

« Parvenir, grâce à des dispositifs numériques spécialement conçus pour cela, à co-produire avec le grand public un ensemble d'éditions structurées de bonne qualité, seul moyen de former un véritable corpus d'étude ; expérimenter de nouvelles formes de médiation numérique ».

Cette évolution a permis l'apparition de pratiques innovantes dans le but de mieux valoriser les fonds. Ces nouvelles pratiques visent aussi une production de qualité et très structurée et qui sera réutilisée avec plus d'aisance. Cette analyse sur l'évolution des pratiques nous a permis d'apercevoir qu'à partir de l'année 2014, il y a eu beaucoup plus de projets qu'avant dans les services d'archives. Par ailleurs, on voit aussi l'apparition de sites web élaborés à partir des données issues de ces projets. Ainsi, les Archives nationales confirment : « les fichiers TEI produits seront valorisés sous la forme d'un site web d'édition numérique, dans lequel les contributeurs seront crédités ». De même que les Archives départementales d'Indre-et-Loire qui prévoit une édition grand public à l'issue de son projet.

Ces créations font partie des souhaits exprimés par les services d'archives et de leurs réalisations en matière de réutilisation des données produites dans le cadre des projets collaboratifs. Cela marque donc une évolution des pratiques qui prennent une forme de plus en plus élaborée pour le bénéfice des publics et qui permet la valorisation des archives. En effet c'est une nouvelle mise en forme de l'information et il en découle une meilleure exploitation des archives. Cette évolution peut être liée à un besoin d'intéresser plus de contributeurs ou de changement de stratégies de valorisation de la part des services d'archives

Au fil des années le travail confié aux utilisateurs devient de plus en plus sérieux en passant par l'édition de documents, à titre d'exemple ; et aujourd'hui la transcription de texte anciens (manuscrits). Le travail demandé à l'internaute dépend du projet mais aussi des fonds en question. Ainsi, la question qu'on se pose est : est-ce-que tous les projets que nous venons

⁶² XML TEI est un standard informatique utilisé notamment pour produire des éditions électroniques de documents manuscrits rigoureuses et faciles à exploiter. On ajoute au texte transcrit des informations sur la structure de ce texte, sur sa mise en page et sur la nature des objets qui le composent.

d'identifier sont proposés au hasard aux publics ? Le choix des projets est-il réfléchi en fonction des fonds ? Y a-t-il un impact sur la valorisation ?

2. Quels projets pour valoriser quels fonds ?

Après l'analyse que nous avons faite sur les différents types de projets, nous constatons de grandes tendances qui se dessinent au niveau de la pratique du *crowdsourcing* dans les services d'archives. C'est en fonction des fonds que les services d'archives mettent en place tel ou tel projet. Ainsi, il faut savoir choisir le projet le mieux adapté au fonds concerné. Cela est très important pour la valorisation. Un mauvais choix conduirait à un échec du projet et ne permettrait pas non plus la mise en valeur du fonds. Il en résulte une mauvaise description et cela peut même nuire à l'exploitation. Par ailleurs le bon choix du projet peut être un facteur de réussite. Dans ce cas le projet sera la forme la plus appropriée à valoriser le fonds en question.

Nous constatons donc que l'annotation et l'indexation se font sur les registres paroissiaux, les recensements de population, les registres matricules militaires, les matrices cadastrales, les données nominatives et l'état civil, alors que l'identification s'effectue plutôt sur les fonds iconographiques et cartographiques. L'édition et la transcription sont réservées aux testaments et carnets de guerre. Cela peut s'expliquer par la nature de chacun de ces types de documents précités. Par exemple les testaments et les carnets sont adaptés à la transcription et à l'édition car ce sont des textes manuscrits rédigés par les soldats. Ces manuscrits sont pour la plupart des lettres de correspondance entre les soldats et leurs familles, des testaments rédigés par les combattants avant d'aller en guerre, des carnets de poilus. La transcription fidèle permet ainsi aux internautes de comprendre les états d'âme, les émotions des soldats et de leurs proches ; mais aussi de s'informer sur le vécu quotidien et les événements qui ont eu lieu pendant la guerre. Malgré les difficultés de lecture, les manuscrits peuvent être transcrits et édités sous une forme numérique afin d'en permettre l'exploitation plus aisée par les usagers. Cela participe ainsi à valoriser les écrits des soldats.

De l'autre côté les documents nominatifs (état civil, registres matricules militaires, recensements de population, registres paroissiaux et matrices cadastrales) sont destinés à l'indexation et l'annotation. Ces deux activités permettent de relever les informations que ces documents contiennent : noms, prénoms, dates de naissance, âge, lieux de résidence, numéros de propriétés, etc. De ce fait les informations saisies permettent d'enrichir la description et de catégoriser les documents en fonction des dates, des noms de famille, des lieux, etc. Elles servent aussi à décrire et à donner plus de sens à ces documents qui en étaient dépourvus. Elles

permettent également à rechercher par noms, date, lieux dans les moteurs de recherche. Ces informations permettent ainsi de donner de la valeur aux archives. Par conséquent grâce à ces deux activités que sont l'annotation et l'indexation les documents nominatifs sont valorisés au bénéfice des utilisateurs.

Enfin nous constatons que l'identification s'effectue sur les documents photographiques et les cartographies. Ces deux types de documents sont souvent des images sans légende ni date qui peuvent porter sur des bâtiments, des évènements, des monuments, des situations, etc. Le travail d'identification des internautes, souvent résidents ou connaisseurs de l'histoire des villes et des lieux en question, permet de caractériser les documents et de retrouver les informations utiles à la description. Mais dans ce cas, contrairement à l'indexation, on ne fait pas appel aux mêmes savoirs des internautes. C'est-à-dire qu'ici il faut que les internautes soient des connaisseurs de la ville ou de la région en question, afin de les reconnaître. Avec le travail d'identification, l'exploitation des documents en est plus facile. Nous en concluons que c'est la pratique la plus adaptée à ces types de documents et c'est la plus convenable à valoriser les photographies et les cartographies.

Dans l'enquête, nous identifions 9 services d'archives (Oise, Orléans, Yvelines, Yonne, Haute-Loire, Cher, Saint-Etienne, Pas-de-Calais, Hérault) qui proposent de l'indexation collaborative sur les fonds identifiés dans le graphique suivant. En ce qui concerne les projets d'annotation, nous trouvons 6 services (Oise, Cher, Var, Ministère des Armées, Yvelines, Haute-Loire) qui l'ont mis en place. Ces projets d'annotation sont proposés sur les mêmes fonds que l'indexation. Également il y a 6 services d'archives (Oise, Lozère, Orléans, Haute-Loire, Bordeaux, Saint-Etienne) qui ont fait des projets d'identification sur les documents photographiques et les cartographies. L'édition et la transcription sont mises à disposition par 6 services d'archives (Oise, Orléans, Pas-de-Calais, Indre-et-Loire (x2), Saint-Etienne, Archives Nationales) sur les carnets de guerre et testaments de soldats.

Selon les types de fonds, nous pouvons classer les projets en 3 catégories selon le graphique suivant :



Figure 21 : Classement des projets selon les fonds

Ce classement des projets selon les fonds nous permet de percevoir les relations entre ces deux éléments. Ce graphique nous donne aussi une idée des stratégies de mise en valeur des services d'archives. L'importance de cette stratégie réside dans le choix du projet en fonction du type de fonds. Les services d'archives ne choisissent pas les projets au hasard. Il y a donc une réflexion voire une « étude d'adaptabilité » entre les fonds et les projets à mettre en place. Ce travail est très important pour la valorisation et constitue une condition pour réussir la mise en valeur.

a) Répartition des différents fonds présentés dans les projets

Dans la représentation suivante nous donnons la part de chaque fonds dans les projets identifiés dans l'enquête. Cela nous permet d'identifier les fonds et mieux les citer dans la partie qui suit celle-ci.



Figure 22 : Fonds présentés dans les projets collaboratifs

b) Intérêts des fonds

Les services d'archives choisissent les fonds qu'ils présentent en fonction de l'intérêt qu'ils présentent pour les publics. De ce fait c'est pour ces intérêts qu'ils estiment de les mettre en valeur pour que les publics puissent les exploiter dans les meilleures conditions. Les intérêts des fonds pour les publics sont très divers. Nous pouvons en retenir quelques-uns :

<p>Registres matricules militaires</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Généalogie</p> <p><input type="checkbox"/> Histoire militaire</p>	<p>Registres paroissiaux, état civil, fonds publics de données nominatives</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Histoire</p> <p><input type="checkbox"/> Commémoration</p> <p><input type="checkbox"/> Généalogie</p> <p><input type="checkbox"/> Statistique</p> <p><input type="checkbox"/> Histoire locale</p>
<p>Fonds généalogiques</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Généalogie</p>	<p>Fonds historiques</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Histoire locale</p>
<p>Testament de Poilus</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Intérêt historique</p> <p><input type="checkbox"/> Intérêt mémoriel</p> <p><input type="checkbox"/> Connaissance des poilus de la Grande Guerre</p>	<p>Fonds iconographiques</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Richesse du fonds</p> <p><input type="checkbox"/> Ancienneté du fonds</p>
<p>Fonds cartographiques</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Histoire locale</p>	<p>Carnets de guerre</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Intérêt mémoriel</p>

Fiches des Morts pour la France de la Première Guerre mondiale

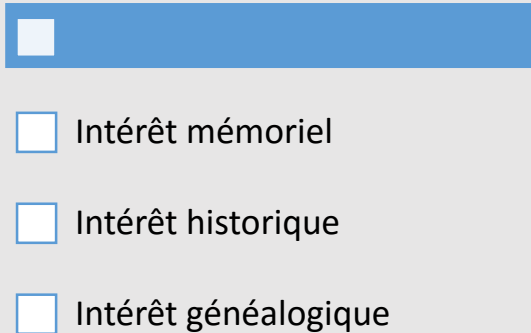


Figure 23 : intérêts des fonds pour les publics et les archivistes

Après ce constat nous pouvons nous poser la question à savoir : est-ce que ces projets sont réellement adaptés aux fonds correspondants ou est-ce que c'est une pratique suivie et copiée mutuellement par les services d'archives au fil des années. Autrement dit est-ce que par exemple les registres matricules militaires sont plus adaptés à l'indexation ou c'est la mode dans les archives ?

Les intérêts des fonds peuvent aussi déterminer les motivations et les objectifs des publics. Ces éléments peuvent être déterminants pour la valorisation des fonds. Car les publics sont des acteurs centraux dans les projets collaboratifs et leurs choix peuvent être influents sur le projet de *crowdsourcing* et sur sa réussite.

3. Les publics des projets : types et motivations

Les publics sont des acteurs importants pour la réussite d'un projet de *crowdsourcing* qui a des retombées considérables sur la valorisation des archives. Les stratégies de mise en valeur des archivistes doivent être en parfaite connivence avec les attentes des publics.

Dans notre enquête, nous avons noté que dans certains projets de *crowdsourcing*, les publics spécialisés sont très motivés que le grand public. Mais il peut arriver aussi dans d'autres projets que le grand public soit très motivé. L'intérêt des fonds pour le public est très important pour la réussite. Il y a des cas où les projets concernent certains fonds spécifiques qui intéressent des spécialistes comme les généalogistes par exemple ou des historiens. Ces types de publics sont très motivés et ont des intérêts spécifiques. Certains publics sont même très présents à certaines activités des services d'archives. Ainsi, les Archives Nationales affirment : « Nous avons aussi

rencontré un petit groupe de possibles contributeurs en 2017 (des personnes qui contribuaient déjà à d'autres activités participatives aux Archives départementales des Yvelines). »

Ces contributeurs sont des collaborateurs, ils ne sont pas de simples participants. Ils ont dépassé le niveau de participant et deviennent des collaborateurs en s'impliquant dans les activités de mise en place et d'organisation des projets auxquels ils vont participer.

Même s'ils ne sont pas toujours présents dans les salles de lecture, certains contributeurs se manifestent aussi dans d'autres projets collaboratifs. Dans ce cas, les Archives nationales affirment toujours :

« Les personnes que nous avons rencontrées en 2017, et certaines qui participent au projet de transcription des testaments, ne sont pas forcément des habitués des salles de lecture, mais ont déjà contribué à des projets participatifs en archives, notamment travaux d'identification de testaments sur place, ou plateforme des AD 78 ».

Cet engagement de la part des contributeurs peut dépendre des types de publics. Nous identifions 5 types de publics dans les services d'archives répondants :

- Grand public
- Associations de généalogie
- Public de Wikisource
- Universitaires
- Généalogistes

Ainsi, 11 services d'archives ont particulièrement ciblé le grand public dans leurs projets. 3 (Archives départementales du Cher, Archives municipales de Saint-Etienne et Archives départementales Var) ont ciblé des publics spécifiques en particulier les généalogistes, les universitaires et associations de généalogie. Les Archives départementales d'Indre-et-Loire, en plus des internautes, a choisi comme cible le public de Wikisource qu'on nomme aussi les wikisourciens. Les Archives départementales de la Haute-Loire ont ciblé par défaut le public de leur site internet. Le public ciblé dans un projet est déterminé par les types de fonds proposés mais aussi par la stratégie de valorisation adoptée par l'organisme. Cette stratégie concerne le projet à mettre en place selon le fonds, comme nous l'avons décrit dans la partie précédente. Tous ces éléments sont d'une importance capitale pour la valorisation. Un autre élément déterminant pour la réussite d'un projet, est la motivation des publics. Ces derniers ont d'ailleurs diverses motivations suivant les projets, les types documents et leurs centres d'intérêts. Tous les services d'archives répondants n'ont pas précisé les motivations de leurs

publics, parfois par manque d'informations sur les usagers. Dans la liste qui suit, nous avons quelques motivations des publics exprimées par les services d'archives :

- Goût pour l'histoire ;
- Sentiment de partager dans un projet d'intérêt général ;
- Devoir de mémoire ;
- Recherches historiques et généalogiques ;
- Participation à une œuvre collective ;
- Altruisme ;
- Travail collaboratif ;
- Esprit d'entraide ;
- Enrichissement du patrimoine numérique local ;
- Faire partager des connaissances ;
- Intérêt des documents ;
- Envie de faire un geste pour la mémoire et pour l'histoire.

Bien vrai que nous n'avons pas toutes les réponses des services à cette question, nous estimons cette liste quasi suffisante pour statuer sur le caractère désintéressé des publics. En effet ils ne sont pas motivés par l'appât du gain ni par un sentiment de reconnaissance. Néanmoins nous n'excluons pas qu'il peut y avoir des publics intéressés par la reconnaissance ou un désir de légitimation ou par d'autres centres d'intérêts. Mais cette liste montre un engagement fort de la part des publics qui donnent de leur temps sans attendre une quelconque contrepartie. Cependant c'est une collaboration gagnant-gagnant car il en découle de ses projets, une mise en valeur des fonds et une meilleure exploitation des archives pour le bénéfice des usagers.

4. Les moteurs de réussite des projets et de la valorisation des fonds

La réussite des projets collaboratifs induit une meilleure mise en valeur des fonds. Fort de ce constat nous estimons que développer ici les moteurs de réussite des projets, c'est en quelque sorte exposer les piliers de la valorisation des fonds. Cela est valable pour les projets dont les objectifs visés sont la mise en valeur des fonds. Dans ce cas la réussite du projet conduit à l'atteinte des objectifs qui sont la valorisation. N'empêche que même pour les projets dont les objectifs sont autres que la mise en valeur, les pratiques du *crowdsourcing* contribuent à la valorisation des fonds à travers ses impacts. Ainsi, nous développerons les finalités de ces projets dans la toute dernière partie de ce mémoire.

Les critères qui permettent d'évaluer la réussite d'un projet collaboratif sont multiples. Ils peuvent provenir des archivistes et des initiateurs des projets, des publics, des outils et des supports utilisés dans les projets.

a) *Les motivations et objectifs des archivistes*

Les motivations et les objectifs des archivistes sont un moteur principal pour la réussite des projets. Ils sont aussi un fondement solide pour atteindre la valorisation. Dans l'enquête que nous avons menée, 80% des services d'archives (12) visent la valorisation de leurs fonds. Les 20% (3 services) ont avancé d'autres motivations: « faciliter l'accès nominatif à certaines typologies, identifier la localisation de certaines photographies » (Archives départementales de l'Oise), « permettre d'indexer des fonds numérisés » (Archives départementales du Var), « la démarche collaborative » (Archives départementales des Yvelines). Cependant les services qui visent la valorisation de leurs fonds d'archives ont des objectifs qui se rapportent aux autres missions archivistiques. Or dans la partie concernant la valorisation des archives, nous étions parvenus avec Martine Cardin à la conclusion selon laquelle toutes les actions archivistiques peuvent être considérées comme des actes de valorisation. Dans le tableau suivant nous présentons, sans rentrer dans les détails, les différents objectifs et motivations exprimés par les services d'archives. Nous faisons correspondre ces éléments avec les missions archivistiques. Concernant celles-ci nous nous appuyons sur l'ouvrage de Carol Couture⁶³ sur les missions de l'archivistique contemporaine.

Objectifs et motivations	Missions archivistiques correspondantes	Services d'archives
Contribuer à améliorer la description des fonds	La description	Archives Nationales
Instaurer une démarche de co-construction des inventaires		Archives Bordeaux Métropole
<ul style="list-style-type: none"> • Offrir un service innovant et participatif pour fidéliser le public du site • Participation au Grand Mémorial pour les registres matricules 		Archives départementales du Cher

⁶³COUTURE, Carol et collab. *Les fonctions de l'archivistique contemporaine*. Québec : Presses de l'université du Québec, 1999. 559 p. (Gestion de l'information). ISBN 2-7605-0941-9

Créer une communauté d'entraide		Archives départementales de la Haute-Loire
Contribuer à l'indexation de la base du Grand Mémorial	L'indexation	Archives départementales de l'Hérault
<ul style="list-style-type: none"> • Mettre en avant l'intérêt du carnet • Valorisation d'une numérisation 	La valorisation	Archives départementales d'Indre-et-Loire
Mise en valeur des fonds iconographiques	La valorisation	Archives départementales de la Lozère
<ul style="list-style-type: none"> • Permettre une interrogation plus fine de la base la plus consultée du site • Permettre une connaissance historique plus précise 	La diffusion	Ministère des Armées
<ul style="list-style-type: none"> • Faciliter l'accès nominatif à certaines typologies • Identifier la localisation de certaines photographies 		Archives départementales de l'Oise
Faire découvrir les fonds	La valorisation	Archives municipales et communautaires d'Orléans
Valoriser des fonds conservés aux Archives départementales	La valorisation	Archives départementales du Pas-de-Calais
<ul style="list-style-type: none"> • Favoriser l'accès et la connaissance • Améliorer la fréquentation du site 		Archives municipales de Saint-Etienne
Permettre d'indexer les fonds numérisés	L'indexation	Archives départementales du Var
<ul style="list-style-type: none"> • Recherche de mains-d'œuvre • Permettre l'indexation nominative de tous les actes d'état civil 	L'indexation	Archives départementales de l'Yonne

<ul style="list-style-type: none"> • Annoter l'ensemble des fiches matricules des soldats de la Grande Guerre pour offrir des données exploitables pour la recherche • Adopter une démarche collaborative 	L'indexation	Archives départementales des Yvelines
---	--------------	---------------------------------------

Figure 24 : objectifs et motivations exprimés par les services d'archives

Dans ce tableau nous constatons que la majorité des motivations des services peuvent se rapporter aux fonctions archivistiques. Par ailleurs il manque des fonctions qui sont liées aux pratiques du *crowdsourcing* notamment l'acquisition. Cela peut s'expliquer par le fait que le projet d'ajout de documents, considéré comme une forme de collecte d'archives, n'a pas été mis en place par les services d'archives de notre échantillon. Mais si nous nous intéressons aux seules motivations exprimées ici, nous concluons que les services d'archives mettent en place des projets collaboratifs pour la description, l'indexation et encore plus pour la valorisation de leurs fonds. Cette valorisation que visent les services d'archives est la fonction la mieux réalisée par le *crowdsourcing*. Car même si les acteurs ne l'expriment pas, les projets collaboratifs permettent toujours de mettre en valeur les fonds en facilitant l'exploitation par la description la plus fine possible des pièces d'archives.

b) La participation des archivistes eux-mêmes

La participation des archivistes est aussi importante que celle des publics. Les archivistes doivent donc participer autant que les publics à l'enrichissement des fonds dans les projets. Certes ça fait partie de leur métier via la description mais dans les projets collaboratifs où ils font appel au public, ils ne doivent pas laisser tout le travail aux contributeurs. Au contraire ils doivent participer et donner un bon exemple au public. Ainsi, ça peut être une source de motivation pour les publics quand ils voient que les archivistes qui portent le projet sont profondément engagés. Cela peut donc contribuer à la réussite du projet et aider les initiateurs à atteindre leurs objectifs. Cette participation peut aussi s'opérer dans la vérification et le contrôle des contributions des publics. Les archivistes peuvent donc jouer le rôle de modérateur dans les propositions des contributeurs. Précisons que ce n'est pas une tâche facile car les archivistes peuvent parfois ne pas détenir l'information qui permet de valider les propositions. Si c'était le cas ils ne feraient pas appel au public. Par contre nous avançons cette idée pour dire qu'après le travail de mise en place des projets, les archivistes ont la possibilité soit de participer

dans l'enrichissement ou de valider, vérifier les propositions s'ils détiennent les informations nécessaires.

c) La participation des publics

La participation des publics est un facteur de réussite majeur des projets. Cette participation est quantitative mais aussi qualitative. Ces deux types de participation sont indispensables pour évaluer un projet collaboratif. Ainsi, nous l'avons reporté dans le questionnaire d'enquête. Les réponses à la question fermée sur les critères d'évaluation des projets peuvent être :

- L'implication et l'engagement des publics : qualité
- Le nombre d'utilisateurs ayant participé : quantité

En effet le nombre d'utilisateurs est un des critères d'évaluation de la participation quantitative. D'autres critères quantitatifs peuvent être identifiés, entre autres le nombre de documents indexés, transcrits par exemple, le nombre de tags déposés, etc. Par ailleurs la participation qualitative peut être notée dans la qualité des contributions des publics. Cela est possible grâce à l'implication du public qui est suscitée par les motivations. Les motivations sont donc importantes pour définir le niveau des contributions car elles peuvent expliquer la bonne ou la mauvaise qualité des données produites par les publics.

Dans les réponses des services d'archives nous avons identifié 54% des réponses sur les critères d'évaluation qui sont en faveur de la qualité contre 46% qui sont pour la quantité. Cela nous donne donc une idée que la réussite d'un projet collaboratif doit être évaluée sur la base des contributions de bonne qualité. Pour qu'il y ait de la qualité sur les propositions, il faut un contrôle. Certains services d'archives ont des procédures d'évaluation de la qualité des données produites. Ainsi, les Archives nationales expliquent :

« Les contributions soumises à validation sont relues par l'équipe projet ; celles dont le niveau de qualité est bon sont validées, les autres sont remises en jeu, avec des commentaires pour explicatifs, pour que les contributeurs les corrigent et complètent. »

De même que les Archives départementales d'Indre-et-Loire déclarent qu'il y a : « relecture et validation des contenus en cours », le Ministère des armées qui nous dit : « vérification et validation par le webmestre », et les Archives Bordeaux Métropole qui avancent : « vérification par les archivistes en charge du fonds », etc. Mais cette évaluation n'est pas effectuée par tous les services d'archives. Dans ce cadre les Archives départementales du Cher déclarent « nous ne l'évaluons pas, nous laissons les internautes se contrôler mutuellement et nous intervenons en juge de paix pour trancher d'éventuels problèmes, mais il n'y en a pas pour ainsi dire ».

Pour que la qualité des données des contributeurs soit élevée, il faut un minimum de vérification et de modération de la part des initiateurs, même si ce n'est pas systématique. Car d'une part les contributions de bonne qualité reflètent au mieux le contenu des fonds d'archives, permettent leur meilleure valorisation des fonds. D'autre part si les données produites par les publics sont de bonne qualité, elles peuvent être réutilisées sous différentes formes et à des fins diverses.

d) La réutilisation et la mise en valeur des contenus produits par les publics

La réutilisation des données produites par les publics peut être faite par différentes manières. Elles peuvent être utilisées pour enrichir les bases des données des services d'archives et être valorisées dans la recherche. Elles peuvent aussi être mises en ligne par certains services d'archives pour que tous les internautes en bénéficient (Archives départementales de l'Oise). Les services d'archives ont donné diverses réponses à la question sur la réutilisation des données produites par les internautes. Nous avons réuni leurs réponses dans un tableau synthétique ci-dessous. Notons que tous les services d'archives n'ont pas répondu convenablement à la question. Dans ce cas nous n'avons reporté que celles qui nous ont semblées adéquates.

Réponse	Service d'archives
Diffusion sur le site internet des Archives et sur le Grand mémorial	Archives départementales de l'Hérault
Enrichissement des analyses	Archives Bordeaux Métropole
Les fichiers TEI produits seront valorisés sous la forme d'un site web d'édition numérique, dans lequel les contributeurs seront crédités	Archives Nationales
Elles sont mises à disposition du public internaute	Archives départementales de la Haute-Loire
Une édition grand public et une restitution à la famille du poilu	Archives départementales d'Indre-et-Loire
Export de données pour les mairies, associations, ou autres organismes sur demande	Ministère des Armées
Elles sont mises en ligne et accessibles à tout internaute	Archives départementales de l'Oise

Elles sont mises en ligne pour recherche	Archives municipales et communautaires d'Orléans
Elles sont mises à disposition des données à l'ensemble des internautes	Archives départementales du Pas-de-Calais
Elles sont mises en ligne	Archives municipales de Saint-Etienne
L'indexation est en ligne est à la disposition des internautes et des lecteurs	Archives départementales de l'Yonne
Ces données sont au service des réutilisateurs	Archives départementales des Yvelines
Mise à jour des analyses	Archives départementales de la Lozère

Figure 25 : réutilisations des données produites par les publics

Ce tableau nous informe que les données des internautes sont produites à des fins utiles. En effet elles sont utilisées pour enrichir les analyses et descriptions des documents, mais aussi réutilisées sous des formes plus valorisantes. Les deux réutilisations qui nous intéressent dans ce tableau sont celles données par les Archives nationales et les Archives départementales d'Indre-et-Loire. Ce dernier service affirme que les la transcription réalisée par le public sert à faire une édition grand public et sera restituée à la famille du poilu. Cela est intéressant dans la mesure où ça peut motiver d'autres familles de soldats qui détiennent des manuscrits de la guerre à les offrir aux services d'archives. C'est donc une excellente mise en valeur des archives. C'est le cas également pour les Archives nationales qui précisent que les fichiers TEI sont valorisés sous la forme d'un site Web d'édition numérique et mieux, les contributeurs seront crédités dans cette œuvre. C'est un puissant argument pour amener d'autres contributeurs et motiver les participants à fournir un travail de qualité.

La réutilisation des données produites par les publics peut donc être un motif pour intéresser d'autres publics. Car elle permet de mettre en valeur le travail des internautes et des fonds en question. C'est aussi permettre l'exploitation des archives dans une forme plus appropriée. Les initiatives prises par les services d'archives à savoir la création de bases, l'édition ou la restitution sous une forme numérique quelconque (sites web) sont très intéressantes. Elles constituent une des meilleures formes de valorisation des fonds.

En plus des éléments que nous avons décrits et qui permettant de valoriser les archives, il y a d'autres éléments qui peuvent participer dans cette valorisation, notamment les supports de la médiation.

e) Les supports de la médiation numérique

Avec l'émergence des outils 2.0, de nouvelles opportunités s'offrent aux archivistes. Ces derniers disposent en effet d'un large panel d'outils pour échanger avec leurs publics, disséminer l'information et mettre en valeur leurs fonds d'archives. Les outils choisis par les services d'archives font partie de la stratégie de valorisation des fonds. En plus ils peuvent être déterminants pour la réussite des projets.

Par supports de médiation numérique, nous entendons les outils qu'utilisent les services d'archives pour réaliser leurs projets collaboratifs. Mais avant de rentrer dans les détails de ces types de supports, faisons d'abord une nette opposition entre les outils existants et les outils développés par les services d'archives dans le cadre des projets collaboratifs. En nous basant sur la littérature et sur les résultats de l'enquête que nous avons menée, les supports de médiation numérique peuvent être répartis en trois catégories : les sites internet des services d'archives, les réseaux sociaux et les plateformes spécialement dédiées aux projets. Ces trois catégories identifiées peuvent aussi être scindées en deux parties : les systèmes existants et ceux développés par les services d'archives.

Pour faire du Web collaboratif, les services d'archives utilisent pour la plupart des systèmes existants (plateformes de réseau social et site internet institutionnel).

o *Les sites web des institutions / services d'archives*

Beaucoup services d'archives utilisent leur site web institutionnel pour faire du web collaboratif. Édouard Bouyé, en 2012, avait présenté 21 services d'archives qui ont développés eux-mêmes des modules sur leurs sites web. Dans notre étude, il y a 13 services d'archives qui affirment avoir utilisé leur site web pour réaliser leurs projets collaboratifs. Ces services sont : Bordeaux, Cher, Haute-Loire, Hérault, Lozère, Ministère des Armées, Oise, Orléans, Pas-de-Calais, Saint-Etienne, Var, Yonne, Yvelines. Dans ce cas le travail se fait de connivence avec des opérateurs comme Arkothèque qui a eu 2 clients (Ministère des armées, Saint-Etienne). Outre cet opérateur, le marché du Web collaboratif a d'autres sociétés notamment Libresens et Tequilarapido, Archinoé, Ligeo, Mnesys [Bouyé, 2012, p. 130].

Ces plateformes développées par les services sur leur site institutionnel peuvent être attractif et intéressant pour les publics. Car elles donnent un aspect nouveau aux sites. Elles peuvent aussi amener plus de contributeurs dans les projets. Cependant si ces contributeurs sont des habitués du site, ces services peuvent conduire à un manque de repères pour eux ; à moins qu'ils ne soient que des modules qui viennent s'ajouter au reste du site. Il faut donc que les services d'archives travaillent de concert avec les usagers et les opérateurs de façon à ce que ces services ne puissent pas être un inconvénient sur la participation des publics.

- [Les réseaux sociaux](#)

Certains services publics d'archives font recours aux réseaux sociaux numériques pour faire la médiation de leurs fonds. En fait c'est une bonne méthode pour être à la rencontre des usagers et prendre un peu de leurs temps passé sur le web.

Par ailleurs ces usages s'intègrent le plus souvent dans une stratégie dont les réseaux sociaux ne sont qu'une fine partie. Les réseaux sociaux contribuent donc à mettre en place ou à compléter une politique de *crowdsourcing*.

Dix services d'archives utilisaient Flickr et autres systèmes analogues pour l'identification d'images en 2012 [Édouard Bouyé, p. 128]. De même que le projet Photos Normandie, qui concerne les archives de la Bataille de Normandie, a été mis en place sur Flickr. Les autres réseaux sociaux comme Facebook, Twitter peuvent aussi être utilisés, non pas pour initier des pratiques collaboratives, mais aussi pour diffuser des articles, témoignages, bulletins, etc. C'est donc un bon moyen de vulgarisation des archives.

Outre les réseaux sociaux, les services d'archives utilisent d'autres plateformes collaboratives comme Wikisource. Dans notre étude, certains projets de transcription et d'édition de carnets ont lieu sur Wikisource à l'instar des Archives Départementales d'Indre-et-Loire à l'occasion du Centenaire de la Grande Guerre. Notons aussi que l'utilisation des réseaux sociaux et autres plateformes existantes est bénéfique pour les services d'archives parce que c'est à moindre coût. Néanmoins le développement de nouveaux outils, même s'il coûte cher aux services d'archives, est plus innovant et montre leurs engagements et leurs investissements.

- [Les plateformes dédiées](#)

À côté de ces systèmes existants, les services d'archives développent aussi de nouvelles plateformes spécialement dédiées aux pratiques collaboratives. Ces systèmes sont en effet

dissociés aux sites internet institutionnels. Les Archives nationales, les Archives départementales d'Indre-et-Loire et les Archives départementales de l'Oise nous déclarent avoir développé des plateformes dédiées à leurs projets.

En guise de conclusion nous pouvons dire que les outils utilisés par les services, qu'ils soient des systèmes existants ou à développer sont importants pour la réussite des projets. Ils sont aussi de véritables moyens de dissémination et de valorisation des fonds d'archives sur le Web.

5. Les finalités des projets de *crowdsourcing* sur les fonds

Au contraire d'autres domaines culturels, la politique de diffusion sur le web des services d'archives est largement centrée sur la mise en valeur des documents. De ce fait les projets de *crowdsourcing* ont beaucoup d'impacts sur les fonds, entre autres : augmenter la visibilité des documents, leur donner de la valeur et permettre leur meilleure exploitation.

a) *Une valeur ajoutée sur les fonds*

La mise à disposition des usagers de documents numérisés, images voire textes OCRisés, permet des opérations de *crowdsourcing* ambitieuses qui enrichissent considérablement la description des archives : indexation, identification de photographies, correction d'OCR ou transcription collaborative. En effet, les projets collaboratifs apportent une valeur ajoutée non négligeable à la description des documents. Cet enrichissement par les internautes apporte de nouveaux éléments aux archives. Ces nouveaux éléments sont des informations utiles aux archivistes et leur permettent de compléter leur travail. Les services d'archives eux-mêmes proclament cette valeur ajoutée que procurent les projets collaboratifs aux fonds d'archives. Dans ce cadre, les Archives départementales du Cher attestent une « valeur ajoutée des annotations » sur les fonds et les Archives Bordeaux Métropole qui déclarent que les projets apportent d' « utiles compléments d'information ». Cette plus-value permet aux documents d'archives d'être visibles sur le Web.

La valeur ajoutée qu'apportent les projets réside aussi dans la qualité des données. Cela est justifiée par la réutilisation des données que font les services d'archives. Elle est faite sous diverses formes : édition grand public ; mise à disposition ; création de documents, de bases données ; etc.

b) Visibilité des fonds

En plus de cette valeur ajoutée, les pratiques collaboratives procurent une visibilité aux archives. De fait, les informations ajoutées par les internautes permettent aux documents d'apparaître dans les résultats de recherche car ils sont décrits avec des métadonnées sociales. Ce sont donc « des informations qui correspondent davantage aux besoins et usages des internautes : besoins de recherche en plein texte, granularité de description plus fine, bases de données nominatives, géolocalisation des documents, etc. » [Moirez, Moreux, Josse, 2013, p. 5]. Ainsi, ces informations valorisent les documents dans la recherche et également sur le Web. De ce fait les Archives départementales du Cher certifient que les projets apportent une « visibilité des fonds numérisés ». Le ministère des armées ajoute que les annotations « permettent une interrogation plus fine de la base la plus consultée du site (1,4 millions de noms) et une connaissance historique plus précise ». Il précise également que ces pratiques permettent d'inscrire le site dans le web 2.0. En plus, cet enrichissement permet d'accéder facilement aux documents. C'est fort de ce constat que les Archives départementales de l'Oise avancent que les projets permettent « un accès facilité aux recensements de population » et en conséquence aux bases de données gérées par les services d'archives.

Par le fait d'un accès et d'une recherche facilités, les pratiques du *crowdsourcing* permettent de faire découvrir les fonds à un public plus large.

Les pratiques du *crowdsourcing* permettent également à des archives privées conservées par des familles d'être connues du grand public. En effet des pratiques collaboratives comme l'ajout de documents proposé par les services d'archives, donnent une visibilité à des archives (cartes postales, photographies ou archives personnelles) qui étaient jusque-là dissimulés voire inconnues. Grâce aux opérations de *crowdsourcing*, ces archives sont exploitées, valorisées au grand plaisir des familles et des chercheurs intéressés.

c) Meilleure exploitation des archives

Par ailleurs il en résulte une exploitation plus aisée des archives. La description collaborative permet aux internautes de profiter davantage du contenu des archives. Cela est possible car la description n'est pas soumise à un contrôle de langage, elle se fait avec celui des usagers. Par conséquent les archivistes utilisent les données produites par les publics pour enrichir et mettre à jour les analyses. Ainsi, le *crowdsourcing* permet une meilleure connaissance des fonds et une appropriation des documents *in fine* par les usagers. Cette meilleure exploitation des

archives qui découle de ces pratiques collaboratives est très importante à souligner, c'est un bénéfice incommensurable pour les archivistes et pour les usagers. Et cette exploitation figure même parmi les objectifs des projets de *crowdsourcing*, faciliter l'exploitation des documents d'archives.

Aujourd'hui, les éditions et l'encodage en TEI XML des données produites par les internautes est aussi une nouvelle et meilleure exploitation des archives. Par le fait de la transcription et de l'édition des manuscrits, les archives deviennent plus lisibles et facilement exploitables.

Toutefois, la coexistence de données produites par des professionnels et de données produites par les internautes nécessite d'apporter une grande vigilance à la qualité des données produites. [Moirez, Moreux, Josse, 2013, p. 5]

Conclusion

Le *crowdsourcing*, reposant sur l'intelligence collective, permet aux services archives de faire appel aux internautes et bénéficier d'une description de leurs fonds en quantités importantes et à moindres coûts. Le *crowdsourcing* « c'est un volume extraordinaire de contributions, le regroupement de compétences particulières, qui permet de faire appel aux compétences scientifiques et à l'expertise des chercheurs ». (Moirez, Moreux, Josse, 2013]

Outre les gains économiques les archivistes cherchent aussi la valorisation de leurs fonds, qui est une mission centrale dans leur métier. La valorisation fait partie des fonctions de l'archivistique contemporaine et constitue une mission à part entière des établissements et des services d'archives. Elle concerne toutes les archives à travers des objectifs différenciés : la valorisation des documents d'administration et celle des archives patrimoniales à travers une médiation culturelle⁶⁴.

Cette médiation est aujourd'hui facilitée par l'évolution du web avec le phénomène 2.0 qui marque l'apparition d'un nouveau paradigme de communication. Le Web 2.0 place ainsi l'utilisateur au centre de l'Internet et devient un espace collaboratif.

Les pratiques de *crowdsourcing* favorisées par cette nouvelle génération du Web, facilitent la valorisation des archives par la réutilisation des données produites par les publics. En termes de valorisation des archives par les pratiques du *crowdsourcing*, c'est l'édition qui semble mieux le faire avec la réutilisation des données produits par les publics. Les réutilisations les plus intéressantes qui ont été notées, ont été réalisées à partir de projets d'édition. Et c'est l'édition qui permet mieux de valoriser le travail des contributeurs et par conséquent les archives. L'édition n'est-elle donc pas le projet le plus adapté à la valorisation des archives ?

⁶⁴ Françoise Hiraux, *Opt. Cit.*, p. 5

Sources

- Portail National des Archives (*Archives de France*)
 - Centenaire de la Grande Guerre :
<<https://francearchives.fr/en/article/88482585>>
 - Pratiques collaboratives : les internautes acteurs aux côtés des services d'archives : <<https://francearchives.fr/article/26288084>>

Bibliographie

Web 2.0

1. Cardon Dominique, Perrier Valérie Jeanne, Le Cam Florence *et al.*, « Présentation », *Réseaux* [en ligne], 2006/3 (n° 137), p. 9-25. Disponible sur : <https://www.cairn.info/revue-reseaux1-2006-3-page-9.htm> (consulté le 18/05/2018)
2. CHAIMBAULT, Thomas. *Le Web 2.0 : l'avenir du Web*. Dossier documentaire. Enssib, 2007. Disponible sur : <http://www.enssib.fr/bibliotheque-numerique/documents/2-web-2-0-l-avenir-du-web.pdf> (consulté le 17/12/2017)
3. DURIEUX, Valérie « Collaborative tagging et folksonomies. L'organisation du web par les internautes », *Les Cahiers du numérique* 2010/1 (Vol. 6), p. 69-80.
4. O'REILLY, Tim. What is Web 2.0? *O'Reilly Media* [en ligne]. 2005. Disponible sur : <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html> (Consulté le 7/12/2017).

Intelligence collective et archives

1. ARNAUD, Béatrice, CARUSO CAHN, Sylvie. La boîte à outils de l'intelligence collective. Malakoff : Dunod, 2016, 192 p.
2. BOUHAÏ, Nasreddine, SZONIECKY, Samuel, dir. Intelligence collective et archives numériques : vers des écosystèmes de connaissances. London : ISTE Éditions, 2017. 238 p.
3. BOUYE, Édouard. Le Web collaboratif dans les services d'archives publiques : un pari sur l'intelligence et la motivation des publics. *La Gazette des archives* [en ligne]. 2012, n°227, pp. 125-136. Disponible sur : http://www.persee.fr/doc/gazar_0016-5522_2012_num_227_3_4974 (consulté le 02/01/2018)
4. FERBER, J. *Les systèmes multi-agents : vers une intelligence collective, informatique, intelligence artificielle*. Paris : InterEditions, 1997.
5. GASSIE, Brigitte, MOREL, Alexandre, PICHON, Aude, et al. L'intelligence collective dans un contexte de développement professionnel continu : ACoPé, l'exemple d'une communauté de pratique. In : CNAM. *Biennale internationale de l'éducation, de la formation et des pratiques professionnelles, Juin 2015, Paris*, Paris : CNAM, 2015, 15 p. Disponible sur : <https://hal.archives-ouvertes.fr/hal-01188175> (Consulté le 02/01/2018)

6. HASSOUN, Mohamed, EL HACHANI, Mabrouka. *Intelligence collective et organisation des connaissances : 7^e colloque international du chapitre Français de l'ISKO, Lyon, 24-26 juin 2009*. Lyon : Université Jean Moulin Lyon 3, 2009, 399 p.
7. LAMY, Florence, MORAL, Michel. *Les outils de l'intelligence collective : la favoriser, la comprendre, la stimuler*. Malakoff : InterEditions, 2013, 240 p.
8. LEVY, Pierre. *La sphère sémantique : Tome 1, Computation, cognition, économie de l'information*. Paris : Hermès Science-Lavoisier, 2011, 415 p.
9. MAYERE, A. « *L'intelligence collective : une notion en chantier* », *Intelligence collective Rencontre 2006*. Paris : Presses de MINES, 2006
10. PRIE, Y. Vers une phénoménologie des inscriptions numériques. Dynamiques de l'activité et des structures informationnelles dans les systèmes d'interprétations (HDR). Lyon : Université Claude Bernard-Lyon 1. 2011.
11. SUROWIECKI, James, RIOT, Elen (trad.). *La sagesse des foules = The Wisdom of crowds*. Paris : Éditions Jean-Claude Lattès, 2008, 384 p. ISBN : 978-2-7096-2891-4
12. VIDAL, Jérôme. *Jérôme Vidal, le blog de... Jérôme Vidal : articles, entretiens, traductions et autres trucs*. [En ligne]. (Publié en 2010). Disponible sur : <http://jeromevidal.blogspot.com/2010/07/la-question-de-lagency-puissance-et.html> (Consulté le 08/07/2018)

Contribution en ligne, participation et médias sociaux

1. AUBERT, Aurélie ; DENOUEËL, Julie ; GRANJON, Fabien. *Médias numériques et participation : entre engagement citoyen et production de soi*. Paris : Éditions mare & Martin, 2014, 208 p. ISBN : 978-2849341391
2. GARCIA, José Luis, HEATON, Lorna, PROULX, Serge, dir. *La contribution en ligne : pratiques participatives à l'ère du capitalisme informationnel*. Québec : Presses de l'Université du Québec, 2014, 256 p.
3. Proulx, Serge. La puissance d'agir d'une culture de la contribution face à l'emprise d'un capitalisme informationnel. Premières réflexions. *Journal du Mauss* [en ligne], 2011. Disponible sur : <http://www.journaldumauss.net/./?La-puissance-d-agir-d-une-culture> (consulté le 09/03/2018)

Crowdsourcing

1. BOUHAÏ, Nasreddine, SZONIECKY, Samuel, dir. *Intelligence collective et archives numériques : vers des écosystèmes de connaissances*. London : ISTE Editions, 2017. 238 p.
2. BURGER-HELMCHEN, Thierry, « Crowdsourcing : définition, enjeux, typologie », *Management & Avenir* 2011/1 (n° 41), p. 254-269
3. ESTELLES-Arolas E., GONZALEZ-Ladron-De-Guavara F., « Towards an integrated crowdsourcing definition », *Journal of Information Science*, vol. 38, n°2, 2012.
4. FRANÇOIS, Mathilde. *Redocumentarisation et crowdsourcing : les enjeux du Web collaboratif* [en ligne]. Mémoire de Master, archives. Angers : Université Angers, 2016, 135 p. Disponible sur : <<http://dune.univ-angers.fr/fichiers/14003330/20162MHD5507/fichier/5507F.pdf>> (Consulté le 02.01.2018)
5. *Gamification*. El gamificator [en ligne]. Disponible sur : <<http://www.elgamificator.com/gamification/gamification-definition>> (consulté le 09/03/2018)
6. JOSSE, Isabelle ; MOIREZ, Pauline ; MOREUX, Jean Philippe. *État de l'art en matière de crowdsourcing dans les bibliothèques numériques*. Rapport L-4.3.1 du projet de R&D du Fond Unique Interministériel 12. Paris : FUI [en ligne], 2013, 77 p. Disponible sur : <http://www.bnf.fr/documents/crowdsourcing_rapport.pdf> (Consulté le 03/01/2018)
7. MOIREZ, P. Archives participatives. In : AMAR, Muriel, MESGUISH, Véronique, dir. *Bibliothèques 2.0 à l'ère des médias sociaux*. Paris : Editions du Cercle de la Librairie, 2012, 217 p. (Collection Bibliothèques). ISBN : 978-2765413400
8. NEROULIDIS, Ariane. *Le crowdsourcing appliqué aux archives numériques : concepts, pratiques et enjeux*. Mémoire de Master recherche, archives numériques. Villeurbanne : ENSSIB, 2015, 109 p.
9. Pratiques collaboratives : les internautes acteurs aux côtés des services d'archives. francearchives.fr [en ligne], 2018. Disponible sur : <<https://francearchives.fr/article/26288084>> (consulté le 05/04/2018)
10. Production participative. *JORF*. 5 août 2014, n°0179, texte n° 91. Légifrance [en ligne]. Disponible sur : <<https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=3E3F01A4E2A178BD068>>

[7960539544F0F.tplgfr25s_3?cidTexte=JORFTEXT000029331922&dateTexte=&oldAction=rechJO&categorieLien=id&idJO=JORFCONT000029330499](http://www.archivesnext.com/?p=2319)>Consulté le 10 décembre 2017)

11. THEIMER, K., *Exploring the Participatory Archives*, 2011, Disponible sur : <http://www.archivesnext.com/?p=2319> (consulté le 21-10-2017).

Valorisation des archives

1. Association des archivistes du Québec. L'exploitation et la mise en valeur des archives à l'ère numérique. *Archives*, Volume 45, Numéro 1, 2013-2014, 256 p.
2. Association des archivistes Français. *Les mutations du métier d'archiviste et de son environnement* : Actes du 11^{ème} colloque national des archivistes communaux et internationaux du 2 au 4 juin 2015 à Limoges. La Gazette des Archives. 2016/4, n°244, 254 p.
3. HIRAUX, Françoise, MIRGUET, Françoise. *La valorisation des archives. Une mission, des motivations, des modalités, des collaborateurs. Enjeux et pratiques actuels*. Louvain-la-Neuve : L'Harmattan-Academia, 2012. ISBN : 978-28061-0059-7
4. PIPON, Brigitte ; LAUBIE, Xavier. Module 12-Valorisation des archives. PIAF [en ligne]. Disponible sur : www.piaf-archives.org/se-former/module-12-valorisation-des-archives (consulté le 01/04/2018).
5. COUTURE, Carol et collab. *Les fonctions de l'archivistique contemporaine*. Québec : Presses de l'université du Québec, 1999. 559 p. (Gestion de l'information). ISBN 2-7605-0941-9

Dictionnaires et lexiques

1. Centre National de Ressources Textuelles et Lexicales. Définition du verbe « actualiser ». Disponible sur : <http://www.cnrtl.fr/lexicographie/actualiser> (Consulter le 12/05/2018).
2. Dictionnaire terminologique archivistique : Disponible sur : https://francearchives.fr/file/4575c619ab1e1e738d81d2249ff8dd4115a3d8cb/ARCHIVES_DE_FRANCE_Dictionnaire_de_terminologie_archivistique.pdf (consulté le 02/04/2018)

Table des annexes

Annexe 1 : Questionnaire d'enquête

Annexe 2 : exemple réponse de services d'archives

Enquête sur les pratiques de crowdsourcing/projets collaboratifs

Cette enquête vise à étudier le crowdsourcing sous l'angle de la valorisation des archives numériques publiques dans le cadre d'un mémoire de recherche.

Veillez identifier votre institution/service d'archives. *

Réponse courte

Avez-vous mis en place un ou des projets de crowdsourcing au sein de votre structure/service d'archives ? *

Réponse longue

Quels types/nature de fonds s'agit-il dans ce (s) projet (s) ? *

Réponse longue

Quel est l'intérêt de ce (s) fonds proposé (s) ? *

Réponse longue

En quoi consiste le (s) projet (s) en général ? *

- Indexation, tagging
- Relecture
- Annotation
- Edition (carnets)

Identification d'images, de lieux

Ajout de photos, cartes postales

Transcription

Depuis quand avez-vous mis en place ce (s) projet (s) ? Combien de temps ont-il duré ? Si votre réponse concerne plus d'un projet, merci de préciser la durée de chaque projet. *

Réponse longue

Aviez-vous ciblé un public en particulier au départ ? Si oui, merci d'identifier les types de public concernés. *

Réponse longue

Est-ce que ceux qui mettent en place ce (s) projets (s) connaissent les contributeurs ? Si oui, de quelles façons ? *

Réponse longue

Les contributeurs sont-ils des habitués des salles de lecture, du site web de l'institution ? Se manifestent-ils ailleurs sur le web ? *

Réponse longue

Qu'est-ce qui les motive à participer à votre projet ? Participent-ils à d'autres projets ? *

Réponse longue

Avez-vous utilisé votre site web ou une plateforme dédiée ? *

Avez-vous utilisé votre site web ou une plateforme dédiée ? *

Site web de l'institution

Plateforme dédiée

⋮
Votre institution dispose-t-elle d'un compte sur les réseaux sociaux ?

Cases à cocher

Facebook

×

Twitter

×

Flickr

×

Instagram

×

Autre...

×

Si oui, avez-vous utilisé ce compte pour communiquer sur le (s) projet (s) ? *

Réponse longue

Comment jugez-vous la participation du public ? *

Satisfaisante

Non satisfaisante

Autre...

Comment justifiez-vous votre réponse ? *

Réponse longue

Quels seraient les critères d'évaluation de la réussite de votre projet ?
Pouvez-vous ajouter d'autres critères à la liste suivante ? *

L'implication et l'engagement des publics

Le nombre d'utilisateurs ayant participé

Autre...

Quels ont été vos objectifs ? *

Réponse longue

Qu'est-ce qui vous a motivé en mettant en place ce (s) projet (s) ? *

Réponse longue

Aviez-vous l'idée de valoriser vos fonds à travers ce (s) projet (s) ? *

Réponse longue

Avez-vous cherché à connaître les motivations des publics ? Si oui quelles
sont-elles ? *

Réponse longue

Qu'est-ce que les publics vous ont apporté dans le cadre de ce (s) projet (s) ? *

Réponse longue

Comment évaluez-vous la qualité des données produites par les usagers ? *

Réponse longue

Avez-vous mis en place une procédure pour valider ou vérifier les propositions du public ? Si oui, merci de la décrire. *

Réponse longue

Quelle utilisation faites-vous de ces données produites par les internautes ? *

Réponse longue

Que pensez-vous de ces données produites par les publics ? Ont-elles de la valeur face à la description des archivistes ? *

Réponse longue

Annexe 1 : questionnaire d'enquête sur les pratiques du *crowdsourcing* dans les services d'archives

Enquête sur les pratiques de crowdsourcing/projets collaboratifs

Cette enquête vise à étudier le crowdsourcing sous l'angle de la valorisation des archives numériques publiques dans le cadre d'un mémoire de recherche.

Veillez identifier votre institution/service d'archives. *

Archives nationales de France

Avez-vous mis en place un ou des projets de crowdsourcing au sein de votre structure/service d'archives ? *

oui

Quels types/nature de fonds s'agit-il dans ce (s) projet (s) ? *

testaments de guerre de Poilus conservés au sein des minutes des notaires de Paris (Archives nationales) ou des Yvelines (Archives départementales des Yvelines)

Quel est l'intérêt de ce (s) fonds proposé (s) ? *

Grand intérêt historique (informations sur les Poilus eux-mêmes, leur catégorie sociale, leurs biens, leur usage du français, leur état d'esprit...) et intérêt mémoriel

En quoi consiste le (s) projet (s) en général ? *

- Indexation, tagging
- Relecture
- Annotation
- Edition (carnets)
- Identification d'images, de lieux
- Ajout de photos, cartes postales
- Transcription

Depuis quand avez-vous mis en place ce (s) projet (s) ? Combien de temps ont-il duré ? Si votre réponse concerne plus d'un projet, merci de préciser la durée de chaque projet. *

Le projet a démarré en 2016. La plateforme est en ligne depuis le 30 janvier 2018
(<https://testaments-de-poilus.huma-num.fr>)

Aviez-vous ciblé un public en particulier au départ ? Si oui, merci d'identifier les types de public concernés. *

Le grand public

Est-ce que ceux qui mettent en place ce (s) projets (s) connaissent les contributeurs ? Si oui, de quelles façons ? *

Nous ne connaissons vraiment les contributeurs que parce qu'ils participent aux rencontres organisées sur le projet. La plateforme mise en place leur demande uniquement une adresse de messagerie et un identifiant. Nous avons aussi rencontré un petit groupe de possibles contributeurs en 2017 (des personnes qui contribuaient déjà à d'autres activités participatives aux Archives départementales des Yvelines).

https://docs.google.com/forms/d/1t6ofP9-74AICK2mRWixc55FFySTAD8epfsvgh0UrSI/edit#response=ACYDBNix9m11B8Ax_TmDz162i4LXlmgG3TYRSF8Sqq

12/05/2018

Enquête sur les pratiques de crowdsourcing/projets collaboratifs

Les contributeurs sont-ils des habitués des salles de lecture, du site web de l'institution ? Se manifestent-ils ailleurs sur le web ? *

Les personnes que nous avons rencontrées en 2017, et certaines qui participent au projet de transcription des testaments, ne sont pas forcément des habitués des salles de lecture, mais ont déjà contribué à des projets participatifs en archives, notamment travaux d'identification de testaments sur place, ou plateforme des AD 78

Qu'est-ce qui les motive à participer à votre projet ? Participent-ils à d'autres projets ? *

Les personnes rencontrées en 2017 ou le 30 janvier sont motivées par l'intérêt des documents et par l'envie de faire un geste pour la mémoire et pour l'histoire

Avez-vous utilisé votre site web ou une plateforme dédiée ? *

Site web de l'institution

Plateforme dédiée

Votre institution dispose-t-elle d'un compte sur les réseaux sociaux ? *

Facebook

Twitter

Flickr

Instagram

Autre : Pinterest

Si oui, avez-vous utilisé ce compte pour communiquer sur le (s) projet (s) ? *

Oui, mais la communication s'est aussi et surtout faite par une journée d'études et à l'occasion d'ateliers (qui vont être renouvelés), sans parler de la plateforme web elle-même

https://docs.google.com/forms/d/1t6offP9-74A1ck2mRWixc55FFy5TAD8epfsvgH0UrSI/edit#response=ACYDBNix9m11B8Ax_TmDz162i4LXimgG3TYRSF8Sqqq

Comment jugez-vous la participation du public ? *

- Satisfaisante
- Non satisfaisante
- Autre :

Comment justifiez-vous votre réponse ? *

Grand nombre de contributeurs inscrits en quelques semaines (environ 230 à ce jour); forte activité sur la plateforme ; très nombreux commentaires publiés par ailleurs sur la plateforme par les contributeurs, montrant une grande attention au travail demandé ; niveau de qualité des contributions qui augmente

.....

Quels seraient les critères d'évaluation de la réussite de votre projet ? Pouvez-vous ajouter d'autres critères à la liste suivante ? *

- L'implication et l'engagement des publics
- Le nombre d'utilisateurs ayant participé
- Autre :
La production d'éditions structurées en TEI, de suffisamment bonne qualité pour en permettre l'exploitation par des chercheurs et la restitution sous la forme d'une édition numérique
-

Quels ont été vos objectifs ? *

Parvenir, grâce à des dispositifs numériques spécialement conçus pour cela, à co-produire avec le grand public un ensemble d'éditions structurées de bonne qualité, seul moyen de former un véritable corpus d'étude ; expérimenter de nouvelles formes de médiation numérique

.....

Qu'est-ce qui vous a motivé en mettant en place ce (s) projet (s) ? *

Même réponse que ci-dessus

https://docs.google.com/forms/d/1t6oFP9-74A/Ck2mRvWxc55FFy5TAD8eptsvgH0UrSi/edit#response=ACYDBNbx9m11B8Ax_TmDz162I4LXimgG3TYR8SF8Sqqc

2/05/2018

Enquête sur les pratiques de crowdsourcing/projets collaboratifs

Aviez-vous l'idée de valoriser vos fonds à travers ce (s) projet (s) ? *

Oui bien sûr. Ces testaments sont inédits. Mais au-delà il s'agissait d'en permettre l'étude scientifique et la restitution intelligente au grand public.

Avez-vous cherché à connaître les motivations des publics ? Si oui quelles sont-elles ? *

Voir ci-dessus. Nous connaissons encore mal les contributeurs.

Qu'est-ce que les publics vous ont apporté dans le cadre de ce (s) projet (s) ? *

Il est trop tôt, à tout juste 5 semaines du lancement de la plateforme, pour faire un bilan.

Comment évaluez-vous la qualité des données produites par les usagers ? *

Les contributions doivent respecter les règles de transcription et d'encodage élaborées pour le projet et publiées sur la plateforme. Voir aussi la charte de contribution.

Avez-vous mis en place une procédure pour valider ou vérifier les propositions du public ? Si oui, merci de la décrire. *

Les contributions soumises à validation sont relues par l'équipe projet ; celles dont le niveau de qualité est bon sont validées, les autres sont remises en jeu, avec des commentaires pour explicatifs, pour que les contributeurs les corrigent et complètent. Etc. Le tout directement via la plateforme.

Quelle utilisation faites-vous de ces données produites par les internautes ? *

Les fichiers TEI produits seront valorisés sous la forme d'un site web d'édition numérique, dans lequel les contributeurs seront crédités. Voir la charte du projet : <https://testaments-de-poilus.huma-num.fr/api/web/uploads/ProjetTestamentsDePoilus-charte-version-1.pdf>

Annexe 2 : Réponse des Archives nationales à l'enquête

Entretien

Entretien téléphonique avec Jean-Yves Leclerc, archiviste, conservateur du patrimoine / chef du service des nouveaux usages des archives et gestion à la direction des archives et du patrimoine d'Ille-et-Vilaine

Table des illustrations

Figure 1 : Origine historique et chronologique du <i>crowdsourcing</i>	pp. 20-22
Figure 2 : Représentation des idéologies marxistes et socialistes.....	p. 23
Figure 3 : Représentation des idées libertaires et anarchistes.....	p. 24
Figure 4 : Représentation des idées humanistes.....	p. 24
Figure 5 : Représentation du libéralisme.....	p. 25
Figure 6 : Services d'archives et institutions enquêtés.....	pp. 35-36
Figure 7 : Cartographie des services d'archives de l'échantillon.....	p. 37
Figure 8 : page du projet d'identification d'images d'Indre-et-Loire.....	p. 42
Figure 9 : page d'identification de photographie de la Haute-Garonne.....	p. 43
Figure 10 : espace personnel de l'annotation collaborative d'Orléans.....	p. 45
Figure 11 : espace personnel MyArchive pour l'annotation collaborative de l'Aube.....	p. 46
Figure 12 : page d'indexation des Archives de l'Ain.....	p. 48
Figure 13 : Projet d'indexation collaborative du Ministère des Armées.....	p. 49
Figure 14 : page projet d'indexation collaborative des ponts et chaussées des Archives Nationales de France.....	p. 51
Figure 15 : page de recherche de testaments du projet de transcription collaborative.....	p. 53
Figure 16 : page d'accueil du projet d'édition des Archives du Pas-de-Calais.....	p. 54
Figure 17 : Projets collaboratifs dans les services d'archives répondants.....	p. 55
Figure 18 : Classement des projets de <i>crowdsourcing</i> selon le niveau de difficulté du travail de l'internaute.....	p. 57
Figure 19 : Classement chronologique des projets des services d'archives répondants.....	pp. 58-59
Figure 20 : Chronologie des projets mis en place par les services d'archives répondants....	p. 60
Figure 21 : Classement des projets selon les fonds.....	p. 77
Figure 22 : Fonds présentés dans les projets collaboratifs.....	p. 78

Figure 23 : Intérêts des fonds pour les publics et les archivistes.....pp. 79-80
Figure 24 : objectifs et motivations exprimés par les services d'archives.....pp. 83-85
Figure 25 : réutilisations des données produites par les publics.....pp. 85-88

Table des matières

Sommaire	5
Introduction	7
I. ÉLÉMENTS DE TERMINOLOGIE	9
A. L'intelligence collective	9
1. L'intelligence collective à travers l'analogie de la fourmilière	9
2. La sagesse des foules	12
3. La puissance d'agir des sujet sociaux	14
B. Le contributeur et la contribution en ligne.....	15
C. Le <i>crowdsourcing</i>	18
1. Définitions diverses, sens convergents	18
2. Participation involontaire des internautes.....	19
3. Origines historiques et conceptuelles du <i>crowdsourcing</i>	20
4. Typologie du <i>crowdsourcing</i>	25
D. Autres notions liées au <i>crowdsourcing</i>	26
1. Le Web 2.0	26
2. L'indexation sociale (métadonnées sociales) ou <i>folksonomies</i> ou indexation communautaire	28
3. Les « archives participatives »	29
II. CROWDSOURCING DANS LES SERVICES D'ARCHIVES	32
A. Enquête auprès des services d'archives ayant mis en place des projets de <i>crowdsourcing</i> : méthodologie adoptée	32
1. Conception des outils d'enquête	33
2. Présentation de la population et de l'échantillon de l'enquête.....	33
3. Les services d'archives et institutions répondants.....	36
B. Panorama du <i>crowdsourcing</i> dans les services d'archives.....	39
1. Ajout de documents, de photographies, de cartes postales.....	40
2. Identification de documents iconographiques	41
3. L'annotation collaborative	44
4. Indexation collaborative	46
5. Transcription collaborative	51
6. Edition collaborative de carnets de guerre.....	53
7. Catégorisation des projets	54
III. LA VALORISATION DES ARCHIVES ET LE CROWDSOURCING	62
A. Qu'est-ce que la valorisation des archives ?.....	63

1. La valorisation dans la chaîne archivistique	65
2. Diffuser et valoriser : quelles différences ?	67
3. Forces des services d'archives : masse documentaire, documents uniques et aux valeurs multiples, etc.	69
B. La valorisation des archives par le <i>crowdsourcing</i> : comment ?.....	71
1. Évolution des pratiques du <i>crowdsourcing</i> : quels impacts sur la mise en valeur des archives ?.....	73
2. Quels projets pour valoriser quels fonds ?	75
3. Les publics des projets : types et motivations	80
4. Les moteurs de réussite des projets et de la valorisation des fonds	82
5. Les finalités des projets de <i>crowdsourcing</i> sur les fonds	91
Conclusion	94
Sources	95
Bibliographie	96
Table des annexes.....	100
Entretien	113
Table des illustrations.....	114
Table des matières.....	116