



HAL
open science

Indexation collaborative et participation sur le web : l'exemple du Mobilier national

Caroline Bogue

► **To cite this version:**

Caroline Bogue. Indexation collaborative et participation sur le web : l'exemple du Mobilier national. Art et histoire de l'art. 2019. dumas-04395795

HAL Id: dumas-04395795

<https://dumas.ccsd.cnrs.fr/dumas-04395795>

Submitted on 15 Jan 2024

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Caroline Bogue

Indexation collaborative et participation sur le web : L'exemple du Mobilier national.

Mémoire de stage
(2^e année de 2^e cycle)
Documentation et humanités numériques

Présenté sous la direction de Mme Françoise Dalex
& M. Christophe Leclercq

Maître de stage : Mme Hélène Cavalié

Septembre 2019

Le contenu de ce mémoire est publié sous la licence *Creative Commons*

CC BY NC ND



AVANT-PROPOS

Ce mémoire trouve son inspiration dans l'expérience professionnelle menée lors de ma deuxième année de 2^e Cycle de l'École du Louvre intitulée « *Documentation et Humanités numériques* », le sujet ayant été proposé par Hélène Cavalié, conservateur du patrimoine chargée du service de documentation du Mobilier national, et approuvé par Françoise Dalex et Christophe Leclercq, responsables de ce Master. Il vient souligner le mouvement d'ouverture qu'initie actuellement le Mobilier national pour offrir à ses interlocuteurs et ses publics des espaces de contacts privilégiés avec les collections et ses responsables. Pour cela, le *crowdsourcing* est apparu comme une possibilité à explorer. Ce terme anglo-saxon correspond à une pratique qui touche un très large éventail de champs d'action, de sujets et d'individus car son principe fondateur repose sur un phénomène viral, Internet, et sur des valeurs partagées massivement, l'échange, le devoir de transmission des objets et des idées, la recherche incessante d'un mieux-être et d'un lien à l'Autre et à l'Ailleurs.

C'est justement par intérêt pour cette capacité qu'ont les hommes à se mobiliser autour de corpus de données, d'objets constitutifs de notre quotidien et de notre patrimoine, et de problématiques citoyennes, que j'ai décidé de consacrer le temps qui m'était alloué pour investiguer les tenants et les aboutissants de ces formes d'engagements. Descriptions, transcriptions, indexations, annotations, enrichissements, créations de contenus.... les possibilités d'actions sont nombreuses et les modalités pour les rendre aisées, intéressantes, incitatives, parfois même addictives pour certaines personnes capables d'y consacrer une grande partie de leurs loisirs, sont très instructives lorsque nous les examinons et essayons de déduire les solutions les plus adaptées au contexte du Mobilier national, détenteur de collections matérielles exceptionnelles et de savoir-faire qui se doivent d'être préservés en un moment de notre histoire où « *l'intelligence de la main* » fait l'objet d'une attention toute particulière¹.

Afin de prendre la mesure des enjeux de la participation sur le web, j'ai visité de nombreux projets collaboratifs en ligne, testé moi-même certaines plateformes afin de comprendre les

¹ Créé en 1999, le prix Liliane Bettencourt pour l'intelligence de la main® récompense savoir-faire, créativité et innovation dans le domaine des métiers d'art. Il est devenu un label d'excellence des métiers d'art français et contribue à leur rayonnement.

interactions qui se nouent entre les ressources, les outils pour les exploiter et les contributeurs qui vont les utiliser, et de saisir pourquoi certains projets se maintiennent dans le temps et d'autres pas. Ces expériences se sont aussi enrichies de la lecture d'articles de recherches et de thèses qui foisonnent ces dernières années à mesure que le recours au crowdsourcing et aux sciences citoyennes tend à devenir une méthode de recherche à part entière. En outre, je me suis attachée après avoir saisi les particularités du Mobilier national, de ses collections et des objectifs poursuivis ces deux dernières années pour mettre à disposition ses collections et les connaissances qui y sont liées, à interroger des professionnels ayant eux-mêmes mis en œuvre des projets collaboratifs. Je souhaitais en effet recueillir leur vécu, enregistrer les points d'attention cruciaux pour la gestion de telles initiatives, demander des conseils à propos des logiciels, de l'hébergement et autres aspects techniques de la question. L'idée étant ensuite de pouvoir établir les modalités d'actions pour les concepteurs qui choisiront de nous aider bénévolement à l'amélioration du rayonnement du patrimoine culturel dont le Mobilier national est le détenteur.

Concevoir un projet participatif demande une réflexion sur le fond comme sur la forme très poussée, comprenant la définition d'objectifs clairs auxquels les professionnels de l'institution comme le public pourront adhérer. D'où le temps nécessaire pour penser chaque brique de conception, élaborer le modèle de données, les interfaces, les modalités de validation, les fonctionnalités et la communication à initier pour susciter la participation et la maintenir sur le long terme. Ce travail est complexe car il faut être conscient des réalités internes et externes à l'institution et évaluer de fait ce qu'il est possible de mettre en œuvre concrètement à courte et moyenne échéance en fonction des moyens qui nous sont octroyés. Au terme de ce stage, les jalons essentiels sont posés mais la mise en œuvre effective demandant plusieurs mois de développement, ce n'est que plus tard, d'ici la fin de l'année, que nous aurons un retour sur ce que nous avons construit et que nous pourrons évaluer les réussites et corriger les failles.

Afin de rendre la lecture de ce mémoire plus fluide, voici l'explicitation de quelques conventions d'écriture adoptées : sigle « MN » sera utilisé pour faire référence au Mobilier national et l'acronyme GLAMs fera référence aux « Galleries, Libraries, Archives and Museums » pour s'accorder avec la terminologie en vigueur dans le monde anglo-saxon.

Pour la réalisation de ce travail, j'ai pu compter sur des personnes très impliquées tout au long de l'année que je tiens ici à remercier. Tout d'abord, mes deux directeurs de mémoire, Mme

Françoise Dalex, responsable du centre de ressources Dominique Vivant Denon, et M. Christophe Leclercq, chercheur associé au médialab de Sciences Po, qui m'ont prodigué conseils et orientations avisés me permettant de mener à bien le projet au cœur de mon stage de fin d'études. J'ai bien sûr une pensée particulière pour Mme Hélène Cavalié, Conservateur du patrimoine en charge du service de documentation du Mobilier national, qui m'a accueillie et aidée tout au long de cette expérience au MN à prendre la mesure des spécificités de cette institution et des attentes de celle-ci en termes de mise à disposition et de valorisation de ses collections sur le web. Et un grand merci à Laurie Chapotte, designer de service, et Ned Baldessin, développeur, les deux membres de l'équipe EIG qui ont mené un superbe travail pour la création du catalogue des collections du MN, qui m'ont prodigué moult conseils pour naviguer avec aisance dans leur univers.

Mes remerciements vont aussi vers toutes les personnes qui ont accepté de m'accorder du temps pour partager leur expérience et leurs ressources au sujet de la création de plateformes collaboratives pour la mise en valeur et la documentation d'objets culturels, de fonds d'archives, de savoirs et savoir-faire :

- Mme Sophie Resseguier, cheffe de la section « technologies des systèmes d'information » au ministère des Armées et M. Christophe Dupont, webmestre du site *Mémoire des Hommes*, pour l'interface d'indexation collaborative *Morts pour la France de la Première Guerre mondiale* ;
- M. Pascal Bertrand, professeur d'Histoire de l'Art moderne à l'université de Bordeaux Montaigne, et Mme Charissa Bremer-David, conservateur du département des sculptures et des arts décoratifs du J. Paul Getty Museum pour *l'inventaire participatif des tapisseries du XVIIIe siècle fabriquées à la manufacture de Beauvais* ;
- M. Jean-Philippe Moreux, expert scientifique Gallica (BnF), pour la plateforme *Correct* ;
- M. Guillaume Raschia, maître de Conférences au Laboratoire des Sciences du Numérique de Nantes, porteur du projet *RECITAL* et responsable du projet *3Puma* ;
- Mme Marie-Françoise Limon-Bonnet, conservateur général du patrimoine aux Archives nationales, responsable du Minutier central des notaires, Mme Emmanuelle De Champs, Professeure d'Histoire et Civilisation britannique à l'Université de Cergy-Pontoise, Mme Florence Clavaud, conservateur en chef du patrimoine, responsable de la mission référentiels et de projets numériques aux Archives nationales, M. Karl Pineau, doctorant en sciences de l'information et de la communication à l'université des technologies de

Compiègne et M. Romain Dugast, chef de service au public des Archives départementales des Yvelines pour le projet de transcription collaborative *Testaments de Poilus* .

- M. Yves Carlier, conservateur général du département des ressources documentaires du château de Versailles, Mme Marine Masure-Vetter, coordinatrice du récolement des collections et Mme Géraldine Bidault, chargée d'études documentaires, photothèque et mise en ligne des collections au château de Versailles pour la création du portail des collections du domaine et château de Versailles.

Enfin, je remercie mes proches pour leur soutien tout au long de cette année, tout particulièrement mes parents et mes attentifs relecteurs. Qu'ils trouvent ici l'expression de ma gratitude pour leur patience et leur générosité.

SOMMAIRE

AVANT-PROPOS

SOMMAIRE

INTRODUCTION.....1

PREMIÈRE PARTIE : LE WEB PARTICIPATIF.....4

QUAND L'INTERNAUTE DEVIENT VECTEUR DE TRANSMISSION DE SAVOIRS ET DE LA MÉMOIRE COLLECTIVE

1. MOBILISER LES « FOULES » POUR ACCROÎTRE LA CONNAISSANCE : PRÉMICES, CONCEPTS, ENJEUX.....4

1.1 LE CROWDSOURCING : LE RENOUVEAU D'UNE PRATIQUE ANCIENNEMENT VALIDÉE.....4

1.2 UN APPEL À LA « SAGESSE » DES FOULES.....8

1.3 VERS UNE REDOCUMENTARISATION DU PATRIMOINE ?.....9

2. CONTRIBUER : POUR QUI, POURQUOI, COMMENT ?.....10

2.1 PARTICIPATION, COLLABORATION, CO-CRÉATION : DÉFINITIONS.....10

2.2 DIFFÉRENTES MODALITÉS DE CONTRIBUTIONS.....11

3. LES MOTEURS DE LA PARTICIPATION.....14

3.1 MOTIVATIONS INTRINSÈQUES ET EXTRINSÈQUES.....14

3.2 VERS L'ÉMERGENCE DE NOUVEAUX ESPACES DE DÉBAT ET DE LIBERTÉ ?.....16

3.3 EFFICACITÉ DOCUMENTAIRE ET DEVOIR DE MÉMOIRE : QUEL POSITIONNEMENT DU CROWDSOURCING DANS CES DEUX PARADIGMES ?.....17

DEUXIÈME PARTIE : INITIER UN PROJET PARTICIPATIF.....19

RÉUNIR LES VOLONTÉS ET LES COMPÉTENCES POUR RÉINVENTER LA RELATION DES INDIVIDUS AU PATRIMOINE

1. LES ENJEUX ÉTHIQUES, JURIDIQUES ET ÉCONOMIQUES AUTOUR DU CROWDSOURCING POUR LES INSTITUTIONS CULTURELLES.....19

1.1 OUVRIR SES COLLECTIONS, SES DONNÉES, SES PRATIQUES : POUR QUELLES RAISONS ?.....19

1.2 DÉFINIR LES CONTOURS D'UN PROJET : DES RESSOURCES ET DES BESOINS.....20

1.3 LE DEVENIR DES DONNÉES PRODUITES : RÉINTÉGRATION, RÉUTILISATION ET PROTECTION.....	22
2. DES INDIVIDUS DERRIÈRE L'ÉCRAN : LE RAPPORT CONCEPTEUR-CONTRIBUTEUR, PROFESSIONNEL-AMATEUR.....	26
2.1 UNE COMMUNAUTÉ D'INTÉRÊT SUR QUI COMPTER.....	26
2.2 LA PERCEPTION DU CROWDSOURCING PAR LES PROFESSIONNELS.....	29
3. DES DONNÉES, DES COMPÉTENCES, DES OUTILS : COMMENT CONCEVOIR LE MODÈLE DE CROWDSOURCING ADAPTÉ.....	31
3.1 LES ÉTAPES DE LA CONDUITE DE PROJET.....	31
3.2 S'ASSURER DE LA QUALITÉ DES DONNÉES PRODUITES : LA QUESTION CHARNIÈRE DE LA VALIDITÉ SCIENTIFIQUE DU PROJET.....	34
3.3 ESTIMER LA RÉUSSITE D'UN PROJET ET SA VIABILITÉ : SUR QUELS CRITÈRES BASER CETTE ANALYSE ?.....	37
TROISIÈME PARTIE : L'EXEMPLE DU MOBILIER NATIONAL.....	39
LE PARTICIPATIF EN GESTATION	
1. S'OUVRIRE AU PLUS GRAND NOMBRE : UN DÉFI À LA MESURE D'UNE GRANDE INSTITUTION.....	39
1.1 L'ENTREPRENARIAT D'INTÉRÊT GÉNÉRAL : UN TREMPLIN.....	39
1.2 CONCERNER LE PLUS GRAND NOMBRE : COMMENT ET POURQUOI IMPLIQUER LE PUBLIC DANS LA VIE DE L'INSTITUTION ?.....	41
2. CRÉER UN PROJET COLLABORATIF : LA RÉALITÉ DU TERRAIN.....	43
2.1 UN PROJET AU LONG COURS ET DES MODALITÉS DE CONTRIBUTION ÉVOLUTIVES.....	43
2.2 DES CONTRAINTES HUMAINES ET FINANCIÈRES À LA MISE EN ŒUVRE DU MODULE PARTICIPATIF : CE QU'IL ADVIENT AUJOURD'HUI.....	46
3. DE LA PROSPECTION À LA CONCRÉTISATION D'UN MODULE DE TRANSCRIPTION COLLABORATIVE : DES DÉBUTS PROMETTEURS.....	49
3.1 UNE COMMUNAUTÉ PROFESSIONNELLE PORTEUSE DES VALEURS DU CROWDSOURCING.....	49
3.2 NOUER DES PARTENARIATS ET LANCER DES EXPÉRIMENTATIONS : LA MANUFACTURE DE BEAUVAIS ET LE PROJET ARACHNÉ.....	53
3.3 LES MODALITÉS DE MISE EN ŒUVRE DU PROJET DE TRANSCRIPTION : PRÉSENT ET AVENIR.....	55
CONCLUSION.....	59
BIBLIOGRAPHIE.....	61

INTRODUCTION

Avec l'avènement du Web 2.0 (aussi appelé Web participatif)² au début des années 2000, l'internaute est peu à peu devenu non plus seulement un usager, récepteur de contenus mis en ligne par des autorités compétentes, mais lui-même créateur de données et acteur incontournable d'un réseau dont le principe repose sur la participation de toutes et tous à la construction du savoir. Des initiatives en ce sens sont très populaires, telle Wikipédia dont l'objectif est de permettre à toute personne détentrice de connaissances, d'écrire de manière communautaire des articles pour une encyclopédie rassemblant tout ce que l'être humain a découvert, créé et transmis au fil des siècles. Les musées ne pouvaient pas rester en marge de ce mouvement, d'autant plus lorsque le cœur de leurs missions (définies par l'ICOM) est l'ouverture au public, la conservation, la diffusion, l'exposition et la transmission du patrimoine culturel matériel et immatériel de l'humanité et de son environnement. Le Mobilier national³, l'une des institutions les plus anciennes en France dédiée à la conservation et la valorisation des savoir-faire au service de sa mission d'ameublement des résidences d'État, se devait d'investir les possibilités offertes par le Web 2.0 et les technologies numériques pour affirmer (et confirmer) sa présence et son unicité dans le paysage muséal et culturel français, voire international.

Cet élan s'est concrètement matérialisé lorsque le MN s'est porté candidat au programme Entrepreneur d'Intérêt Général (EIG), créé en 2016 par le Président de la République afin de recruter des talents extérieurs à la fonction publique pour développer des solutions numériques innovantes répondant aux besoins des administrations. Menant une réflexion depuis quelques années déjà sur l'intérêt d'ouvrir ses collections au public par le biais d'un catalogue en ligne, le MN s'est entouré de personnes internes et externes à l'institution pour valider les orientations stratégiques et opérationnelles et donner de la force à

² O'REILLY Tim, BATELLE John; « What is web 2.0? Design Patterns and Business Models for the Next Generation of Software », in *O'Reilly*, publié le 9 septembre 2005, p. 1, [en ligne], consulté le 19 avril 2019, URL: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>.

³ Mobilier national : <http://www.mobiliernational.culture.gouv.fr/>.

sa candidature. En 2017, deux EIG, une designer et un développeur, ont pu être recrutés pour mener ce projet avec l'équipe de l'institution en dix mois, en interne donc, ce qui fait toute la singularité de cette expérience. En effet, la mise en ligne des collections des musées fonctionne souvent par le recours à des prestataires extérieurs qui répondent à un cahier des charges établi par les responsables des collections et de la documentation. Or dans le cas présent, les besoins et les attentes des publics en termes d'accessibilité et de contenus ont pu être analysés au jour le jour grâce à une communication constante des équipes et une adaptation permanente de l'interface construite brique par brique. Aujourd'hui, le résultat est en ligne et le code ouvert pour toute autre institution souhaitant reproduire cette solution⁴. Mais, ceci n'est qu'une étape de la politique actuelle du MN qui a choisi d'impliquer directement le public par la création d'outils d'indexation et de transcription collaboratives visant à fédérer les volontés pour explorer les sources primaires et révéler enfin en détail l'histoire des objets qui ont accompagné les personnages illustres de notre histoire dans leurs intérieurs et mettre en exergue le savoir-faire des artisans d'art actifs en France.

Notre travail a consisté à identifier à la fois la communauté d'intérêt qui serait prête à participer et qui pourrait apporter son regard et son expertise pour garantir au projet toute sa pertinence et son succès en termes d'enrichissements documentaires. Pour cela, il était indispensable de mener une exploration sélective des projets participatifs antérieurs ou en cours de réalisation afin de prendre toute la mesure du phénomène depuis ses prémices et d'identifier les ressorts cognitifs à sa base, les problématiques rencontrées et les réponses à apporter. Les questions relatives aux outils logiciels et aux interfaces à mettre en œuvre, à la communication et à la médiation à instaurer avec le public, aux leviers incitatifs, à la construction et à l'agencement des contenus, à la création et à la fidélisation des contributeurs, aux personnes qui assureront la réalisation et l'animation, ont été abordées car tout ceci doit être partie prenante de la réflexion à mener en amont du projet afin de garantir sa « réussite ».

Il ne suffit pas d'ouvrir ses collections et ses données à la contribution, encore faut-il évaluer dans quel cadre celle-ci va s'effectuer, quelle qualité de redocumentarisation est attendue de cet appel à la « sagesse des foules »⁵ et comment les apports « amateurs » vont se

⁴ Le code source permettant le développement d'une interface de consultation des collections semblable à celle proposée actuellement par le MN est disponible à cette adresse : <https://github.com/entrepreneur-interet-general/gobelins>.

⁵ SUROWIECKI James, *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economics, society and nations*, New York, Doubleday, 2004, 320 p.

concilier avec l'expertise des professionnels de musées. Car les craintes de l'institution par rapport à la participation du public sont réelles, d'où l'importance du dialogue et du recueil d'expériences afin d'intégrer le changement du rapport entre le public et les collections consécutif à l'essor du numérique. Il ne faut pas oublier que le *crowdsourcing* a pour objectif principal d'offrir davantage de portes d'accès à l'information, que ce soit en facilitant la recherche dans les bases de données, en recontextualisant les sources ou en fournissant des aides à l'analyse d'images et de textes.

Dans une première partie, nous reviendrons sur la notion de Web participatif et sur le changement dans les rapports entre l'internaute et l'information qu'il a induit, et surtout sur les facteurs de cette contribution effective du public. Revenir sur ce qui incite l'internaute à collaborer, c'est se replacer dans le cheminement intellectuel qui conduit à percevoir la foule comme ferment de l'amélioration de l'information et pas uniquement comme consommatrice d'une vérité disséminée par des experts reconnus officiellement. La mise en ligne des données offre différentes modalités de participation que nous définirons ici.

Dans une deuxième partie, nous montrerons comment mettre en œuvre un projet participatif, à travers l'analyse critique de plusieurs initiatives qui nous semblaient exemplaires en termes de conduite de projets, de solutions adoptées et de modalités de participation. Les entretiens menés auprès de porteurs de telles initiatives ont permis de dégager des recommandations à destination du MN en vue de proposer une expérience de collaboration et une qualité des données produites optimales.

Pour achever ce mémoire, la troisième partie offrira au lecteur le déroulement de notre réflexion, depuis la prise en compte des spécificités inhérentes au fonctionnement et aux collections du MN jusqu'à la proposition de solutions pouvant faire converger les énergies de chacun à une meilleure valorisation et diffusion du patrimoine conservé. L'instauration d'une participation réelle et active du public à l'enrichissement documentaire des biens culturels s'effectuera selon différentes modalités dont la mise en œuvre et les résultats adviendront dans un temps qui dépasse la durée de notre stage, le retour d'expérience des usagers, tout comme celui des professionnels du MN, étant attendu dans les six à huit mois qui suivront la mise en ligne de l'interface de collaboration. Projet au long cours, la question du temps est une donnée importante sur laquelle nous aurons l'occasion de revenir.

PREMIÈRE PARTIE : LE WEB PARTICIPATIF

Quand l'internaute devient vecteur de transmission de savoirs et de la mémoire collective

1. Mobiliser les « foules » pour accroître la connaissance : prémices, concepts, enjeux.

1.1. *Le crowdsourcing : le renouveau d'une pratique anciennement établie et validée*

Le terme de « *crowdsourcing* » apparaît pour la première fois sous la plume de Jeff Howe en juin 2006 dans le titre d'un article du magazine *Wired*⁶ alors que le World Wide Web prenait de plus en plus d'importance dans les sociétés humaines, tant dans ses dimensions politiques, économiques que sociales. La capacité de l'information à circuler s'accroît considérablement avec l'émergence d'un Internet interactif et des réseaux sociaux, les usagers sont connectés et ont la capacité d'agir sur la manière dont les structures du pouvoir et du savoir s'articulent. Si certains chercheurs en Humanités numériques, tel Tim O'Reilly, considèrent le Web 2.0 comme un terrain favorable à l'essor de cet esprit de collaboration qui caractérise le *crowdsourcing* en donnant les moyens aux internautes d'influer sur le fond (l'information), la forme (les plateformes adaptées à l'objet de recherche) et la description (commentaires, géo-référencement)⁷, cette technologie n'est cependant pas le déclencheur du recours à une communauté non définie de bénévoles pour l'externalisation de certaines tâches. C'est à dessein que le terme de « main d'œuvre » a été utilisé par Jeff Howe pour qualifier ces volontaires car c'est d'abord dans le cadre de l'économie marchande que le terme de *crowdsourcing* s'est imposé pour désigner l'implication d'un grand nombre de personnes dans la réalisation de tâches qui autrement n'auraient jamais pu être réalisées de manière aussi rapide et exhaustive, moyennant rétribution. Le site de micro-travail Amazon Mechanical

⁶ HOWE Jeff, "The Rise of Crowdsourcing.", in *Wired*, 14, (6), 2006.

⁷ LEBRETON Claire, *Bibliothèques, tags et folksonomies. L'indexation des bibliothèques à l'ère sociale*, Mémoire d'étude, mars 2008, 117 p., Paris, Département des ressources documentaires, ENSSIB, sous la direction de Nicolas MORIN, p. 73-74.

Turk (AMT, créé en 2005) illustre ce phénomène, son principe étant de faire analyser ou produire des informations, par le biais de sous-tâches plus ou moins complexes, utiles à son fonctionnement là où la machine et l'intelligence artificielle ne peuvent encore se passer de l'intervention humaine⁸. Nous reviendrons sur ce point plus loin tant les implications éthiques du recours à ce type de sites ne sont pas anodines.

Pour autant, il ne faudrait pas réduire le *crowdsourcing* à l'assouvissement d'intérêts purement mercantiles mais s'interroger aussi sur la modification du rapport entre les producteurs d'un service et ses consommateurs : les internautes ne se perçoivent plus uniquement comme des récepteurs d'un contenu formaté par des autorités reconnues officiellement selon « un modèle de diffusion hiérarchique, unilatéral et stable »⁹ (modèle *top-down* ou vertical, tel celui de l'enseignant envers l'étudiant, du bibliothécaire vis-à-vis du lecteur, ou du Web 1.0) mais souhaitent au contraire devenir eux-mêmes créateurs de cette information, dans une démarche volontaire, active et réciproque (modèle *bottom-up* ou horizontal où professionnels et amateurs sont engagés dans une relation axée sur la coopération et une atténuation des hiérarchies habituelles). Cette « idéologie du partage ou de la contribution » est un modèle que le développement de logiciels libres illustre bien avec un code source qui est modifié à mesure que les besoins informatiques des entreprises, des institutions et des individus évoluent ; les utilisateurs signalent les *bugs* et leurs souhaits à la communauté de développeurs qui améliorent le code, rendent visibles les modifications qu'ils apportent gratuitement, car cette co-création est fondée avant tout sur le plaisir et la passion. Croire que ce mode de fonctionnement a émergé avec le Web 2.0 serait une erreur, le modèle du « *Just in Time* » mis en œuvre par Toyota où la production est sous-tendue par la demande et le surplus, ou encore la publicité et son système de prescription, annonçaient l'essor d'une économie de l'attention et d'un renforcement du lien entre les agents qui collaborent selon la nécessité de l'individu et de la collectivité.

Signalons aussi que, bien avant la numérisation qui a permis l'accès à des corpus de données à tous sans devoir revendiquer un quelconque attachement académique, la Science - et les sciences naturelles particulièrement-, fut d'abord et avant tout une affaire d'amateurs, organisée en réseau (cabinets de curiosités, correspondance, sociétés savantes) où seule la

⁸ HEDGES Mark, DUNN Stuart, *Academic crowdsourcing in the Digital Humanities: Crowds, Communities and Co-production*, Londres, Chandos Publishing, Chandos Information Professional Series, 2017, p. 13.

⁹ ANDRO Mathieu, *Bibliothèques numériques et crowdsourcing : Expérimentations autour de Numalire, projet de numérisation à la demande par crowdfunding*, 2017, 461 p., Thèse de doctorat en sciences de l'information et de la communication, Saint-Denis, Université Paris 8 Vincennes, sous la direction de Imad Saleh, p. 42.

capacité à apporter des connaissances et à dialoguer importait. De premières occurrences de collaborations entre passionnés sur des sujets précis font date comme *l'Oxford English Dictionary* (OED)¹⁰ qui repose sur la volonté des membres de la Société philologique de Londres de remédier à des dictionnaires anglais jugés lacunaires dans leur recension du vocabulaire en usage ou désuet et laissant à craindre une transmission de la langue parcellaire dommageable pour la culture anglophone. Devant l'immensité de la tâche et des attentes qualitatives, les philologues décidèrent à compter de 1857 de faire appel au public pour collecter des preuves d'usage de certains mots afin de créer la cartographie la plus exhaustive de la langue. Environ deux milliers de corrections furent envoyés sous la forme de notes manuscrites à cette société qui procéda ensuite à l'édition du dictionnaire. Cette démarche ancienne n'est pas sans rappeler celle du projet *Scottish Words and Place*¹¹ dont l'objectif est de recenser et préserver les dialectes locaux. Un autre exemple plus tardif évoque cette même force de mobilisation d'une foule à la réalisation de communs : en 1938, alors que les États-Unis connaissent la plus grave crise économique de son histoire (La Grande Dépression), un groupe de mathématiciens et de physiciens qui ne trouvaient pas de travail à la fin de leurs études imagina le premier supercalculateur au monde, bien avant l'invention de l'ordinateur, fondé sur la participation volontaire de plus de 450 chômeurs à l'établissement de tables de calcul de fonctions mathématiques cruciales pour le progrès des sciences et des techniques, le *Mathematical Tables Project*¹². Enfin en 1981, c'est l'écriture collaborative qui est illustrée à travers le guide de voyages *Lonely Planet* où des individus mutualisent leurs expériences personnelles de voyages pour fournir des conseils pratiques afin de visiter dans les meilleures conditions possibles différentes contrées. Cette initiative connaîtra une postérité avérée et illustre un intérêt pour les récits de vie qui se vérifiera et s'amplifiera avec le Web 2.0. Les communautés d'intérêt qui naissent de ces initiatives, et entretiennent des liens plus ou moins étroits avec les instances patrimoniales et universitaires, permettent de construire de manière originale des connaissances et de dépasser finalement les modalités traditionnelles de production de données¹³. Ces critères de définition caractérisent le *crowdsourcing* appliqué aux sciences humaines mais qu'est-ce qui le différencie des « sciences citoyennes » qui connaissent elles aussi un renouveau majeur avec le Web 2.0 ? Quels sont leurs apports

¹⁰ OED: <https://www.oed.com/>.

¹¹ ROEGEL Denis, *A reconstruction of the Mathematical Tables Project's table of natural logarithms* (4volumes, 1941), [Research Report] LORIA, UMR 7503, Université de Lorraine, CNRS, Vandœuvre-lès-Nancy, 2017, 26 p.

¹² ANDRO Mathieu, *op. cit.* n°9, p. 65.

¹³ HEDGES Mark, DUNN Stuart, *op. cit.* n°8, p. 4.

respectifs du point de vue des modalités de participation, de la nature des données produites par les internautes et des thématiques investiguées ?

Le terme de « sciences citoyennes » a été créé par Alan Irwin¹⁴ en 1995 pour désigner la méthode consistant à amasser des connaissances sur l'environnement grâce à des protocoles de collecte de données établis par des scientifiques et mis en œuvre sur le terrain par un grand nombre de personnes, qualifiées d'amateurs. Que l'initiative porte sur l'évaluation de la répartition des espèces animales (*Garden Birdwatch*¹⁵) et végétales, sur l'observation de l'évolution des paysages, et autres problématiques, le principe repose toujours sur la délégation : les volontaires réunissent le matériau d'étude qui sera analysé ensuite par les chercheurs à l'origine de la collecte. À l'image de certains projets de *crowdsourcing*, les sciences citoyennes s'ancrent aussi sur l'attachement au territoire, à l'histoire locale, les participants se sentant concernés et prêts à consacrer un peu de leur temps à des tâches parfois rébarbatives. Pour contrebalancer la répétitivité induite par certains sujets, les projets actuels mettent l'accent sur la composante démocratique de cet effort avec une marge de liberté de plus en plus conséquente laissée aux participants dans la définition des axes de recherche et le calendrier (échéances, rencontres, publications). Une divergence importante entre les deux typologies réside certainement dans le nombre de contributeurs qu'il est nécessaire de mobiliser : si la quantité de données enregistrées est un critère de pertinence pour la constitution de bases de données solides dans le contexte des sciences citoyennes, les projets de *crowdsourcing*, considérés comme les plus achevés du point de vue de la qualité des échanges, du partage de savoirs, de métadonnées reconnues « de qualité scientifique », sont ceux qui ont mobilisé un petit nombre de passionnés qui ont investi énormément de leur personne dans la réalisation de ce « travail ». Soyons toutefois prudents dans cette assertion car nous avons pu constater combien définir une structure commune à tous les programmes de recherche échappe à la généralisation, chacun d'eux s'élaborant en conformité avec la nature des sources primaires à rendre lisibles et visibles, les objectifs définis en amont et le public potentiellement intéressé. La résonance est forte entre *crowdsourcing* et sciences citoyennes, tout comme l'est le *crowdsourcing* avec la « nouvelle muséologie » formalisée au début des années 1970 par Georges Henri Rivière qui a « introduit dans le champ patrimonial un

¹⁴ ALAN Irwin, *Citizen Science: study of people, expertise and sustainable development*, Londres, Psychology Press 1995, 216 p.

¹⁵ Garden BirdWatch : <https://www.bto.org/our-science/projects/gbw>. Annexe I, p. 45.

schéma de communication favorisant la participation des citoyens »¹⁶, redéfinissant de la sorte le partage des rôles entre institutions et publics.

Les principes et théories qui régissent le *crowdsourcing* se sont donc construits sur une temporalité bien plus ample que ne laissait présager un premier état de l'art, les outils informatiques et numériques lui ayant toutefois donné une ampleur et une technicité inédites.

1.2. Un appel à la « sagesse des foules »

Un projet participatif repose sur l'humain pour pallier les défauts de la technologie et à l'insuffisance de moyens financiers pour traiter une masse de données toujours grandissante. Mais ce postulat de départ peut être davantage affiné : c'est à l'intelligence collective, dans son hétérogénéité, sa richesse et sa complémentarité, que le *crowdsourcing* puise sa raison d'être, cette « sagesse des foules » que James Surowiecki théorisa en 2004¹⁷ pour faire contrepoids au préjugé selon lequel la foule est synonyme d'ignorance et véhicule de désinformation¹⁸. Pour cet économiste, la foule, lorsqu'elle se caractérise par une diversité d'opinions et d'histoires personnelles, par une indépendance de jugement et une décentralisation (c'est-à-dire la capacité à représenter un savoir local et très spécifique), est à même de trouver des solutions à des problèmes irrésolus jusque-là, de coordonner ses efforts pour servir un objectif commun et de coopérer malgré des affects différents¹⁹. Mais cette foule ne peut agir concrètement sans être guidée, d'où la nécessité de *leaders*, de coordinateurs prescripteurs d'une méthodologie, pour faire émerger une expertise de cet ensemble d'identités réunies autour d'un sujet. Des initiatives comme *Wikipédia* ne dépendent pas seulement des nouvelles technologies mais s'insèrent dans une évolution culturelle majeure qui encourage chaque individu à partager ses connaissances au bénéfice d'un progrès commun, brouillant alors les rôles et les hiérarchies qui prévalaient antérieurement avec le Web 1.0 ou dans les établissements d'enseignement et de conservation du patrimoine.

Avec le *crowdsourcing*, si certains projets reposent sur la mobilisation d'un nombre très important de contributeurs, d'autres reposent sur l'addition d'idées indépendantes et sur la

¹⁶ CASEMAJOR LOUSTAU Nathalie, « La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? », in *Études de communication*, 36, 2011, p. 3, [en ligne], consulté le 31 mars 2019, URL : <http://journals.openedition.org/edc/2532>.

¹⁷ Cette idée n'est pas récente : Aristote dans les *Politiques* voyait déjà que la « multitude est meilleur juge » des questions politiques et esthétiques car son jugement est nourri des compétences et des connaissances de tous ; Nicolas de Condorcet (1743-1794) parlait aussi de « *Miracle de l'agrégation* » selon la loi des grands nombres.

¹⁸ NEROUULDIS Ariane, *Le Crowdsourcing appliqué aux archives numériques : concepts, pratiques et enjeux*, Juin 2015, 109 p., Mémoire du Master Archives numériques, Lyon, Université de Lyon, sous la direction de Céline GUYON, p. 12.

¹⁹ SUROWIECKI James, *op. cit.* n°5.

diversité cognitive : ce n'est alors pas la production mécanique anonyme de données qui importe mais bien la possibilité d'aller à la rencontre de personnes et de communautés qui s'engagent pour leur patrimoine, leur histoire, et qui le questionnent de manière inédite parce que de cette agrégation d'interprétations distinctes va naître l'exactitude de la prédiction collective²⁰. Une évolution du *crowdsourcing* sera-t-elle le *nichesourcing*, soit la mobilisation de petits groupes d'amateurs devenus experts sur des thématiques très spécifiques, qui mettent à profit leur hétérogénéité de profils, de points de vue et de parcours pour dialoguer, analyser, prendre des décisions et procéder à des choix de manière plus constructive ? Le constat du Rijksmuseum va en ce sens : plutôt qu'une foule entendue dans le sens du nombre, il a besoin pour mieux valoriser ses collections d'individus, retraités, amateurs, autodidactes, peu importe, qui échangent ensemble pour faire émerger des idées, des associations d'œuvres, des pistes de recherche innovantes, en dehors des sentiers battus.

1.3. Vers une redocumentarisation du patrimoine ?

Constater la croissance des données et des ressources disponibles sur le Net est un fait mais les rendre accessibles en est un autre. En effet, pour pouvoir consulter une source, un document numérisé, encore faut-il qu'il soit référencé, organisé dans un schéma de données qui rende possible sa localisation par mots-clés dans un moteur de recherche. Numériser des archives ou des œuvres d'art ne suffit pas à les rendre visibles, il faut procéder à leur redocumentarisation, c'est-à-dire ajouter de nouvelles strates d'information, de métadonnées qui vont décrire le document numérisé. Ce travail est en principe assuré par des professionnels des métiers de la documentation mais le *crowdsourcing* peut venir suppléer utilement ces derniers, des amateurs²¹ venant indexer le contenu des documents pour structurer les systèmes d'accès et leur conférer *in fine* pérennité et légitimité²². Cette légitimité s'instaure par le regard simultané des professionnels qui normalisent l'enregistrement des métadonnées, et des internautes qui, tout en décrivant les objets selon leur vision des choses, vont refléter la pluralité des interprétations et des usages pouvant en être faits dans une optique de classification libre, ouverte et évolutive. Cet enchaînement est

²⁰ ELLIS Sally, « A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship », in *Libri*, 64(1), 2014, p. 3- 4.

²¹ Le terme "d'amateur" n'est pas à prendre dans un sens péjoratif mais doit plutôt servir à qualifier une personne qui n'a pas suivi de cursus universitaire dans le domaine en question mais qui pour autant est intéressée et a appris seule, ou pour un étudiant qui souhaite recourir à d'autres méthodes pour enrichir sa formation initiale.

²² SALAÜN Jean-Michel, « La redocumentarisation, un défi pour les sciences de l'information », in *Études de communication*, n° 30, 2007, p. 15.

perceptible dans le projet des *Herbonautes* initié par le MNHN²³ en 2012 qui a permis, suite à la numérisation de l'Herbier national, d'enrichir avec de nouvelles métadonnées les échantillons collectés durant des siècles par les naturalistes qui voyageaient dans le monde entier, de réorganiser et relier ces millions d'items, et de permettre ainsi une meilleure communication à propos de la richesse des collections aux scientifiques comme aux botanistes en herbe. C'est un autre regard qui devient possible grâce à cette recontextualisation et surtout l'édification d'un rempart contre l'oubli et l'isolement que provoque la délocalisation d'abord puis la dématérialisation ensuite de ces spécimens de leur milieu d'origine et de l'institution qui les abrite. Plus qu'une redocumentarisation du monde, c'est bien à une remédiation des sources à laquelle nous assistons avec le Web 2.0 qui autorise le rassemblement inédit de documents et un changement de cadre d'interprétation²⁴.

2. Contribuer : pour qui, pourquoi, comment ?

2.1. *Participation, collaboration, co-création : définitions.*

La consultation de la littérature afférente au *crowdsourcing* et de projets en ligne révèle une certaine confusion dans l'usage des termes de « participation » et de « collaboration ». Pour autant, ce sont deux actions distinctes : un projet de type « collaboratif » se caractérise par des contributeurs qui collectent des données suivant un protocole établi par des scientifiques mais ils procèdent aussi à leur analyse et à leur diffusion, endossant un statut hybride de concepteur/bénéficiaire des contenus générés. Un projet de type « participatif » induit, lui, une démarcation nette entre ceux qui sont à l'origine du projet et en contrôlent complètement les modalités, et ceux qui vont appliquer ces règles, contribuer pratiquement et accepter la disjonction totale qui règne entre contributeurs et concepteurs²⁵. Quant à la co-création, c'est le public qui suggère la question de recherche à l'origine du projet et avec des professionnels, ils vont travailler ensemble à l'élaboration et à la concrétisation de l'ensemble de la démarche, les scientifiques ayant ici plus un rôle consultatif que prescripteur²⁶.

La mise en œuvre d'un projet « participatif », « collaboratif » ou « co-créatif » procède d'un choix de l'institution émettrice de l'initiative qui optera pour le processus le plus adapté

²³ Les Herbonautes : <http://lesherbonautes.mnhn.fr/>. Annexe I, p. 19.

²⁴ PÉDAUQUE Roger T., *La redocumentarisation du monde*, Toulouse, France, Cepaduès éditions, 2007, p. 18.

²⁵ NEROUULDIS Ariane, *op. cit.* n°18, p. 16-17.

²⁶ OOMEN Johan, AROYO Lora, *Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges*, 5th International Conference on Communities and Technologies, New York, 2011, p. 139, [en ligne], consulté le 03 février 2019, URL: http://www.iisi.de/fileadmin/IISI/upload/2011/p138_oomen.pdf.

en fonction du corpus d'étude, du public qui pourrait participer, des activités proposées, et surtout du temps que l'équipe-projet peut y consacrer. Les projets participatifs ont tendance à développer des activités qui ne demandent pas un gros investissement en termes de vérification et de communication, avec des outils facilement appréhendables (souvent fondés sur une même suite d'actions répétitives) demandant peu de formation à leur usage. À l'inverse, pour les projets collaboratifs, la qualité du lien entre les contributeurs et l'institution est fondamentale car le sujet du projet demande une réflexion, voire une expertise, plus poussée en termes de méthodologie et d'organisation, le dialogue entre tous les partis étant incontournable. Néanmoins, dire que tous les projets participatifs se basent sur des tâches simples à réaliser serait erroné : si l'idée du Denver Community Museum de créer un musée sans personnel, sans budget, sans lieu dédié tout en proposant une exposition par mois est catégorisée comme un projet « participatif », il n'en demeure pas moins que le défi est ambitieux et demande autre chose que la réalisation de tâches mécaniques. Ceci montre que si la création de catégories pour tenter de saisir le phénomène de *crowdsourcing* est un point de départ important, la prise en compte des particularités du projet et de l'institution émettrice est impérative pour définir les ressorts adéquats auxquels recourir pour atteindre l'objectif espéré et ne pas désorienter ni décourager les contributeurs par l'ampleur de la tâche.

2.2. Différentes modalités de contributions

Pour que l'effort demandé aux internautes soit justement calibré en fonction de la nature du projet et des contenus à produire, plusieurs dispositifs de participation et de collaboration peuvent être envisagés, les uns n'étant pas exclusifs des autres, un agencement subtil de tâches pouvant être un facteur attractif (possibilité pour l'internaute de choisir d'agir de la manière qui lui convient le mieux) et interprétatif judicieux quand il s'agit de croiser des approches et des disciplines différentes. Nous proposons ici un panorama rapide²⁷ de ces dispositifs qui ont été classés selon leur orientation « données » ou leur orientation « dialogue », les premiers étant davantage contraints que les seconds dans la liberté d'expression laissée aux contributeurs²⁸. Lisa Chupin, qui a travaillé sur la plateforme des *Herbonautes*, propose même cinq rubriques : les projets « *diffusionnels*, *rédactionnels*,

²⁷ Nous avons choisi en effet de ne pas détailler plus avant chacun de ces dispositifs, la littérature à ce sujet étant riche. Nous insisterons davantage dans la troisième partie sur les procédés retenus par le Mobilier national.

²⁸ CHUPIN Lisa, *Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés*, 2017, 220 p., Thèse de doctorat en Sciences de l'information et de la communication, Paris, CNAM, sous la direction de Manuel Zacklad, p. 112.

contributifs, attentionnels en mode flux et transmédia » pour caractériser chacune de ces formes de crowdsourcing, afin de rendre sensibles les degrés de connectivité entre les contributeurs, les concepteurs et les contenus qu'exige chacune d'elles²⁹.

Pour faire face à la croissance exponentielle des documents en ligne en raison de la normalisation des procédures de numérisation et de partage des images et des enregistrements sonores et vidéos sur les réseaux sociaux, l'indexation collaborative et le *tagging* apparaissent comme incontournables aujourd'hui. Bien que tous deux orientés « données », à la différence de l'indexation collaborative qui vise à dépouiller de manière exhaustive et très normalisée des corpus d'archives, de photographies, d'objets afin de constituer des bases de données structurées et réutilisables pour des programmes de recherche (ou par des entreprises qui peuvent commercialiser les données de généalogie produites bénévolement), le *tagging* et les folksonomies s'effectuent sur un mode simple et intuitif, sans contrainte de saisie des informations par des champs prédéfinis ou des vocabulaires contrôlés. Indexation comme *tagging* offrent la possibilité de retrouver une information rapidement, un lien d'index ne s'interprétant que collectivement par le choix de critères identiques pour tous les liens, les *tags* introduisent quant à eux une sémantique plus libre, moins formelle et apparaissant comme un complément de l'indexation classique des documentalistes³⁰.

L'identification de photographies est une tâche très classique et relève de la même volonté de produire des données descriptives sur des fonds documentaires encore peu connus. L'adjonction de *tags* (ou mots-clés) apparaît d'ailleurs comme une modalité de participation et de diffusion des images tout à fait adaptée, aisée à mettre en place par les petites structures qui utilisent des plateformes de partage telles *Flickr* où il suffit de créer un compte pour déposer des contenus ou les annoter. La Bibliothèque du Congrès (Washington, USA)³¹ fut une des premières institutions patrimoniales à tester cette option mais cette documentarisation de fonds iconographiques peut aller plus loin, à l'image du projet *PhotosNormandie*³² où la description prend la forme de commentaires argumentés, nourris d'intenses discussions à propos des lieux et des personnes qui apparaissent sur les clichés. Nous sommes ici à la frontière entre la production de données brutes et la co-construction de savoirs, limite qui se retrouve aussi dans certains projets de transcription et de correction d'OCR (Optical Character

²⁹ *Idem*, p. 113.

³⁰ LÉVY François, NAZARENKO Adeline, GUISSÉ Abdoulaye, « Annotation, indexation, et parcours de documents numériques. Consulter, explorer et maintenir les textes réglementaires », in *Documents numériques*, Vol. 13, 2010, p. 147-148, [en ligne], consulté le 03 mars 2019, URL : <https://www.cairn.info/revue-document-numerique-2010-3-page-121.htm>.

³¹ The Library of Congress – Flickr: https://www.flickr.com/photos/library_of_congress/. Annexe I, p. 37.

³² PhotosNormandie : <https://www.flickr.com/photos/photosnormandie/>. Annexe I, p. 38.

Recognition). Du fait d'une marge d'erreur encore très importante lors de l'océrisation des documents textuels tels que les journaux ou les manuscrits, le recours à la correction ou à la transcription manuelle demeure un impératif pour disposer d'un mode de recherche en plein texte et de synthèse vocale pour les personnes malvoyantes. De nombreux projets tournent autour de cette activité, comme l'interface de correction *Trove* de la Bibliothèque nationale d'Australie³³ qui s'est érigée en modèle ou *Transcribe Bentham*³⁴ pour le versant transcription qui a su créer une communauté soudée à même de faire face à un corpus de textes difficiles d'accès. La transcription ouvre la possibilité de révéler des documents inédits à un public très large et d'en faciliter l'étude. Le projet *Testaments de Poilus* est représentatif de cet aspect car après une première phase exploratoire pour localiser ces documents mal connus, a succédé une étape de transcription collaborative qui aboutira prochainement à une éditorialisation de ces testaments, c'est-à-dire à la production d'un discours documenté, réflexif, critique, inscrit dans un temps donné. Tout l'intérêt de la démarche réside dans la structuration qui est donnée à des corpus afin de mieux les conserver et les rendre disponibles pour d'autres usages.

En effet, l'ouverture des données n'a de sens que si elle permet un renouvellement des approches et des lectures, le « *mapping* »³⁵ et le géo-référencement étant deux moyens dont s'est emparé le *crowdsourcing* pour procéder à ceci. Le géo-référencement consiste à reprendre d'anciennes cartes et à mettre à jour les coordonnées géographiques avec les outils actuels pour établir des correspondances entre données anciennes et problématiques de recherche contemporaine, selon les normes en vigueur, afin d'établir des correspondances avec des informations obtenues plus récemment. Une nouvelle couche de métadonnées est ainsi ajoutée pour donner une cohérence, une visibilité et une lisibilité à des documents qui autrement ne seraient pas utilisables. Le *mapping* vise, quant à lui, à créer de toute pièce une représentation spatiale d'un ensemble de données qui n'entretient pas de lien originel mais procède d'un assemblage artificiel pour mettre en exergue des phénomènes étudiés aujourd'hui. *Recogito*³⁶ est une plateforme qui met à disposition du chercheur une zone de géolocalisation où il est possible d'identifier les noms de lieux présents dans le texte étudié (par exemple les sites traversés par Homère dans *l'Odyssée*) et de les « mapper » sur un répertoire géographique global afin d'obtenir un visuel qui ouvre d'autres perspectives et

³³ Trove : <http://trove.nla.gov.au/>. Annexe I, p. 33.

³⁴ Transcribe Bentham : http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham.es. Annexe I, p. 27.

³⁵ Les termes de « cartographie participative » ou de « modélisation spatiale des données » peuvent aussi être observés dans les publications francophones ou canadiennes.

³⁶ Recogito : <https://recogito.pelagios.org/>. Annexe I, p. 12.

d'autres interprétations par ce décloisonnement des disciplines ainsi autorisé³⁷.

3. Les moteurs de la participation

3.1. Motivations intrinsèques et extrinsèques

Mener à terme un projet de *crowdsourcing*, c'est s'interroger sur qui contribuera et acceptera de donner du temps, de sa personne, tout en acceptant des règles de participation parfois contraignantes pour une cause qui ne lui bénéficiera peut-être pas directement. Alors pourquoi participons-nous ? Contribuer, c'est un acte personnel car chacun à ses raisons personnelles de participer mais c'est aussi un processus social « *continu, dynamique et permanent, il se trouve dans des pulsions, des intérêts, des motifs singuliers qui font bouger les individus, qui entrent en interaction réciproque les uns avec les autres* »³⁸. C'est ce que la littérature qualifie alternativement de motivations intrinsèques et de motivations extrinsèques.

Par motivations intrinsèques, nous entendons tous les ressorts cognitifs qui vont faire qu'une personne seule devant son écran d'ordinateur va décider de se connecter sur le site et saisir des métadonnées ou partager son expérience. Les internautes choisissent donc des sujets par lesquels ils sont attirés et par la perspective d'acquérir autrement que par la formation initiale de nouvelles connaissances et compétences. De cet intérêt et de ce choix libre naissent le goût du travail « bien fait », ce qui est un atout considérable pour la réussite du projet. Le projet *Biodiversity Heritage Library* témoigne de ceci, les participants soulignent dans leurs témoignages que la description des illustrations leur a permis d'acquérir un regard affiné et un vocabulaire adéquat pour décrire les espèces illustrées, les caractériser et les classer, et d'avoir conscience de s'inscrire dans le processus de recherche évolutif qui est une caractéristique de la taxonomie³⁹. Pareillement pour les participants au projet *Testaments de Poilus* qui sont à présent capables de coder en XML-TEI des documents manuscrits, à tel point que certains ont acquis le statut de vérificateur⁴⁰. Mais ces mêmes personnes ont aussi besoin de savoir que le travail qu'elles effectuent va bénéficier à d'autres, qu'ils soient eux-mêmes bénévoles ou chercheurs. Ce sentiment d'altruisme est crucial dans l'environnement du *crowdsourcing* car il conduit d'une part à faire émerger des « super-contributeurs » qui vont permettre une progression conséquente et régulière du projet et d'autre part, il permet la formation de

³⁷ SIMON Rainer, BARKER Elton, ISAKSEN Leif and DE SOTO Cañamares Pau, « Linking early geospatial documents, one place at a time: annotation of geographic documents with Recogito », in *e-Perimtron*, 10(2), 2015, pp. 49-59.

³⁸ PROULX Serge, GARCIA José Luis, HEATON Lorna, *La contribution en ligne : pratiques participatives à l'ère du capitalisme informationnel*, Québec, Presses de l'Université du Québec, 2014, p. 35.

³⁹ HEDGES Mark, DUNN Stuart, *op. cit.* n°8, p. 63-64.

⁴⁰ Voir entretien du 11 mai 2019 en Annexe II, p. 57.

communautés d'intérêt vectrices d'une expertise, d'une réflexivité et de valeurs qui font advenir des orientations et des solutions inédites en phase avec les besoins et les attentes de la société. Mathieu Andro parle de « *beauté du geste* », d'une « *activité qui est une fin en soi* », réalisée « *de manière passionnée et désintéressée, sans rechercher une reconnaissance ou une récompense qui risqueraient, au contraire, de diminuer la motivation* »⁴¹. Lorsque le contributeur espère obtenir un retour concret sur l'activité qu'il exerce, que ce retour prenne la forme de la reconnaissance de la part des concepteurs ou des contributeurs, de récompenses, d'une rémunération, nous touchons à ce que les auteurs appellent les motivations extrinsèques, qui ne dépendent plus de la satisfaction personnelle éprouvée mais de l'Autre qui juge, évalue, contrôle. Le registre de valeurs change, l'activité n'est plus une fin en soi mais un moyen d'atteindre des objectifs qui ne sont plus tout à fait individuels mais correspondent davantage à ceux édictés par le collectif et donc à une certaine propension au conformisme et à la normalisation des comportements. Être vu, voir son avis accepté, se sentir utile sont des aspects que Francis Quesnel qualifie de raisons psychologiques amenant à participer⁴². Motivations intrinsèques et extrinsèques ne s'excluent pas mais sont au contraire intéressantes à mobiliser dans un projet, elles permettent de toucher un public plus diversifié à même d'apporter une ouverture d'esprit plus grande à la résolution des questions soulevées dans le cadre de l'initiative. C'est l'outil, l'interface, le mode d'appropriation des ressources choisi qui vont inférer sur la construction de ces représentations, le modèle de participation induisant des rôles dans lesquels les usagers pourront s'identifier, y trouver la satisfaction recherchée, tout en ayant à l'esprit que les causes premières amenant une personne à contribuer ne sont jamais figées⁴³. Laure Fabien, une des indexeuses de la base des *Morts pour la France de la Première Guerre mondiale* dit qu'elle a commencé à participer pour des recherches sur ses ancêtres, pour étendre ensuite son champ d'activité à toute la région des Vosges dans un « *souci de devoir de mémoire* » tout en ajoutant qu'elle a aussi trouvé dans l'activité un certain ludisme inattendu, « *un jeu dans la complétude des fiches indexées par communes, noms de familles, dates ou lieux de décès, lorsque l'on parvient à toutes les finir (j'y vois un rapport avec Pokémon Go, attrapez-les tous... indexez-les tous)* »⁴⁴.

⁴¹ ANDRO Mathieu, *op. cit.* n°9, p. 222.

⁴² FRANCIS Elie, QUESNEL Odile, « Indexation collaborative et folksonomie », in *Documentation-Sciences de l'Information*, Vol. 64, 2007/1, p. 62.

⁴³ DUN Stuart, HEDGES Mark, *Crowd-Sourcing Scoping Study. Engaging the Crowd with Humanities Research*, Londres, Centre for e-Research, Department of Digital Humanities, King's College London, Novembre 2012, p. 15, [en ligne], consulté le 15 mars 2019, URL : <https://stuartdunn.files.wordpress.com/2013/04/crowdsourcing-connected-communities.pdf>.

⁴⁴ Morts pour la France de la Première Guerre mondiale : <http://www.memoiredeshommes.sga.defense.gouv.fr/fr/>.

Par ce retour d'expérience, nous voyons que l'émulation et la compétition sont aussi importantes à encourager, la recherche du podium, le lancement de défis (« 1 jour 1 Poilu » pour indexer toutes les fiches de la base des *Morts pour la France* avant le 11 novembre 2018), l'obtention de diplômes ou de badges comme sur le site des *Herbonautes*, restent tout à fait compatibles avec les impératifs de production de données qualitatives⁴⁵.

Pour conclure sur ce point, la plateforme, quels que soient son parti pris et son positionnement (institutionnel ou non institutionnel), se doit de mettre en œuvre les outils les plus adaptés (logiciels, ergonomiques, pédagogiques, communicationnels) à la question de recherche pour faciliter l'adhésion du contributeur, anticiper les freins psychologiques et socioprofessionnels, et enfin porter la communauté d'intérêt vers une efficacité croissante.

3.2. *Vers l'émergence de nouveaux espaces de débats et de liberté ?*

Les données que produisent les internautes via ces plateformes participatives ne possèdent pas un intérêt seulement pour les institutions qui sont émettrices de ces projets, mais répondent aussi et surtout aux besoins des internautes d'accéder librement à des ressources qu'ils vont décrire selon leurs propres référentiels. C'est une relation différente qui s'instaure entre le public et le patrimoine, à présent accompagnée et non plus imposée par les instances prescriptrices que sont les universités ou les GLAMs. Les indexations de type « folksonomies » permettent une agrégation de concepts qui émanent directement des usagers du web qui vont pouvoir organiser les informations selon leurs habitudes de navigation, trouver ce qu'ils cherchent sans avoir besoin de comprendre comment fonctionnent les outils documentaires que manipulent les professionnels (thesaurus, ontologies, standards de langages, certes rigides dans leur conception mais aussi homogènes, généralistes, ce qui n'est pas sans qualité pour celui qui sait les utiliser). C'est une manière de se libérer d'une certaine rigidité académique qui privatise les données brutes pour aller vers une consultation directe de la source primaire⁴⁶. En recourant au vocabulaire des usagers du Web, nous nous mettons en phase avec les évolutions que la langue connaît avec le temps et nous assurons que les mots-clés utilisés pour décrire les documents seront réactualisés au fur et à mesure des transformations linguistiques. Les bibliothèques sont parmi les premières institutions à avoir saisi tout le potentiel de l'indexation collaborative : laisser l'utilisateur décrire librement une ressource, c'est faire preuve de confiance envers son savoir, reconnaître sa capacité de

⁴⁵ CHUPIN Lisa, *op. cit.* n°28, p. 118.

⁴⁶ NOIRET Serge, « La Digital History: histoire et mémoire à la portée de tous », in *Ricerche Storiche*, 1, vol. 41, January-April. 2011, p. 118.

discernement et son aptitude à autoréguler ses positionnements. Fort de cette reconnaissance, l'utilisateur se sent en retour plus impliqué dans la vie de l'institution et a tendance à accomplir sa tâche avec sérieux. L'association de points de vue et de vécus différents permet enfin « *de créer un modèle global du monde qui serait le reflet des connaissances et des intérêts des utilisateurs mondiaux* »⁴⁷, d'orienter les politiques de conservation, de mobiliser les énergies autour de projets ambitieux ou de sortir des thématiques de leur mutisme par manque de moyens et aussi parce que nous méconnaissons le potentiel de cette « foule » à investir le Web 2.0 pour co-construire son futur, l'histoire et les valeurs qu'elle souhaite perpétuer. Pour éviter des propositions de participation trop formalisées ou restrictives dans la capacité d'action, des chercheurs tentent de développer le principe de recherche-action où les catégories de savoir qui préexistent sont redéfinies grâce à des méthodes, des normes et des outils nouveaux, pour atteindre une « herméneutique numérique » et une « transindividuation » qui favoriseront la discussion et la résolution de controverses⁴⁸.

3.3. Efficacité documentaire et devoir de mémoire : quel positionnement du crowdsourcing dans ces deux paradigmes ?

La littérature à ce sujet est moins abondante, les aspects opérationnels et fonctionnels du *crowdsourcing* ayant davantage mobilisé les chercheurs, néanmoins ceux-ci sont de plus en plus attentifs aux mémoires et savoirs qui se transmettent via ces activités collaboratives et participatives. Qu'entendons d'ailleurs par « mémoire » dans le cadre du crowdsourcing et quelle légitimité reconnaître à cette manière particulière de construire des connaissances en dehors des circuits traditionnels qui ne fait pas consensus quand il s'agit d'évaluer la valeur scientifique des données générées par des « non professionnels » ?

Par mémoire, nous entendons le phénomène social qui structure les relations entre les hommes et permet aux sociétés de se perpétuer dans le temps. Elle n'existe que si nous échangeons les uns avec les autres, racontons notre histoire et écoutons celle des autres, organisons sa transmission et son énonciation tout en intériorisant son caractère éminemment fluctuant, évolutif, à mesure que les événements s'enchaînent, se connectent, s'explicitent les uns les autres⁴⁹. Ces processus de relecture produisent des alternatives à l'immédiateté qui

⁴⁷ DURIEUX Valérie, « Collaborative tagging et folksonomies. L'organisation du web par les internautes », in *Les Cahiers du numérique*, vol. 6, n° 1, 2010, p. 75.

⁴⁸ PUIG Vincent, « Pratiques contributives à l'âge des données », in CARAYOL Valérie, MORANDI Franc (éd.), *Le tournant numériques des sciences humaines et sociales*, Publication de la Maison des Sciences de l'Homme d'Aquitaine, 2015, p. 82.

⁴⁹ SALAÜN Jean-Michel, *op. cit.* n°22, p. 16.

caractérise nos sociétés actuelles, les innovations technologiques depuis les années 1960 et l'avènement du Web 2.0 ayant modifié notre rapport au temps et à l'espace. Offrir la possibilité à chacun de s'arrêter sur des documents qui témoignent de la pensée et des usages d'une époque, leur donner les compétences d'investiguer des corpus d'objets et de sources qu'autrement ils n'auraient pas osé approcher, c'est permettre à tous de changer leur façon de percevoir les choses et de devenir des acteurs des modalités de transmission et de compréhension de ce patrimoine dont ils sont les héritiers. Pouvons-nous alors parler « *d'efficacité documentaire* »⁵⁰ pour juger les apports du *crowdsourcing* qui n'est finalement qu'une méthode d'analyse parmi tant d'autres, pas moins légitime, pas moins faussée si nous nous reportons à la théorie de James Surowiecki sur la sagesse des foules ? Le *crowdsourcing* est une méthode de documentation, un espace privilégié de partage de connaissances, en aucun cas il ne s'érige comme prescripteur d'une vision mémorielle qui se surimposerait à d'autres. Il laisse place au développement de nouveaux schémas narratifs qui doivent être examinés avec attention et prudence, d'une part parce que l'information numérique reste par essence instable et transitoire et d'autre part, parce que « *dans le cadre des applications du Web 2.0, [nous assistons] à la décentralisation plus forte encore de "l'autorité", et à la fragmentation ultérieure de la notion d'auteur* »⁵¹. L'introduction du numérique dans les sciences humaines ne change pas le cœur des métiers d'historien ou de documentaliste : authentifier, contextualiser, critiquer les sources restent des pratiques primordiales. Le devoir de ces professionnels est cependant de mettre leur expertise en phase avec le potentiel inédit d'appropriation documentaire du patrimoine qu'autorise le World Wide Web, et donc d'utiliser les nouvelles technologies pour faire évoluer les relations entre les publics et les institutions culturelles qui ne doivent plus se contenter de simplement faire de la diffusion pour favoriser au contraire l'inclusion et l'appropriation. Enfin, le *crowdsourcing* pourrait remédier à des lacunes en termes de structuration de l'information en ligne qui s'est faite sans recourir aux langages documentaires traditionnels des professionnels de l'information-documentation et qui pourtant sont tout à fait adaptés et applicables à l'échelle du Web après avoir démontré toute leur pertinence et leur efficacité au niveau des GLAMs⁵².

⁵⁰ CASEMAJOR LOUSTAU Nathalie, *op. cit.* n°16, p. 6-7

⁵¹ NOIRET Serge, *op.cit.* n°46, p. 120.

⁵² MENON, Bruno, « Les langages documentaires. Un panorama, quelques remarques critiques et un essai de bilan », in LEBRETON Claire, *Bibliothèques, tags et folksonomies. L'indexation des bibliothèques à l'ère sociale*, mars 2008, 117 p., Mémoire d'étude, Paris, Département des ressources documentaires, ENSSIB, sous la direction de Nicolas Morin, p. 26-27.

DEUXIÈME PARTIE : INITIER UN PROJET PARTICIPATIF

Réunir les volontés et les compétences pour réinventer la relation des individus au patrimoine

De plus en plus fréquemment, les institutions culturelles, et les musées en particulier, incitent le public à interagir avec leurs collections. Mettre en œuvre un programme participatif demande une véritable réflexion de la part des professionnels, tant sur les objectifs poursuivis que sur leur capacité à inclure l'avis « d'amateurs » dans leurs pratiques de recherche. Le recueil d'expériences mené dans le cadre de notre stage au MN nous permet de revenir sur les points cruciaux de la création et la conduite de ce type d'initiatives.

1. Les enjeux éthiques, juridiques et économiques autour du *crowdsourcing* pour les institutions culturelles

1.1. *Ouvrir ses collections, ses données, ses pratiques : pour quelles raisons ?*

À cette interrogation, plusieurs réponses peuvent être amenées. D'une part, et suite à la publication du rapport de la mission « musées du XXI^e siècle » en mars 2017, les professionnels du secteur ont soulevé cette question fondamentale : « *comment peuvent-ils [les musées], au lieu de se présenter comme les seuls 'dépositaires' de l'autorité culturelle et de la mémoire sociale, se convertir en acteurs ou facilitateurs dans la promotion et la conservation des mémoires que fait naître (de façon non linéaire) notre quotidien ?* »⁵³. Cette mission, créée dans le contexte sensible de la vague d'attentats que connut la France en 2015, a pour vocation de faire en sorte que le musée demeure en phase avec les dynamiques contemporaines qui animent nos sociétés et se constitue comme le promoteur d'une chaîne de dialogue entre « experts » et publics, sans condescendance aucune, afin que le savoir et la maîtrise de l'Histoire ne soient pas accaparés par certaines catégories de personnes, avec les

⁵³ EIDELMAN Jacqueline (dir.), *Rapport de la mission « Musées du XXI^e siècle »*, Paris, Volume 1, Synthèse, février 2017, 173 p.

risques de dérives que ce type de comportements comporte. Depuis 1997, le musée McCord d'histoire sociale de Montréal⁵⁴ a fait le pari d'utiliser le Web pour faciliter l'accès à ses collections et développer une pédagogie innovante d'apprentissage de l'histoire, et notamment de l'histoire coloniale. Identification de fonds photographiques, collecte d'objets, co-curation d'expositions avec des établissements scolaires et des communautés, les axes du participatif et du collaboratif ont été particulièrement explorés, l'implication du public étant à la hauteur de l'audace des professionnels de l'institution dans le choix qu'ils firent de s'ouvrir à une « économie de partage » où l'outil numérique apparaît comme une « valeur ajoutée ».

D'autre part, si les objectifs poursuivis sont clairs du côté des institutions comme des usagers, l'organisation de cette collaboration ne doit pas se faire au détriment de la qualité des contenus produits. Chaque élément du projet, depuis la validation et la réintégration des données jusqu'à l'adéquation des activités proposées avec les résultats escomptés, doivent être réfléchis et soupesés pour nous assurer de la fiabilité de la démarche. N'oublions pas enfin que le volontariat demeure historiquement un recours acté et vital pour les musées, les centres d'archives ou les bibliothèques disposant de peu de moyens afin qu'ils puissent améliorer leurs services et justifier les orientations de leur politique culturelle. Ce constat soulève d'ailleurs des controverses au sujet de la non rémunération d'un travail qui en théorie devrait être réalisé par des professionnels dans le cadre légal prévu pour cela⁵⁵.

1.2. Définir les contours du projet : des ressources et des besoins

Quels que soient les objectifs poursuivis par une institution - accroître la visibilité des collections, mettre en valeur les savoirs et les savoir-faire, donner à voir une autre image et s'insérer dans l'écosystème des humanités numériques -, créer un module participatif demande de la part des professionnels de la documentation de préciser leurs attentes en termes de documentation et les bénéfices que pourront en retirer le public comme l'institution. Cibler les corpus de ressources pouvant être exploités dans un cadre participatif, les caractériser (type, volume) pour créer l'activité la plus à même d'en révéler la richesse informative (transcription, annotation, suggestion de sources), évaluer la possibilité de segmenter le processus en plusieurs tâches pour démultiplier les accès et permettre aux contributeurs de trouver le mode d'intervention qui leur convient le mieux, sont des éléments à définir en

⁵⁴ Musée McCord : <https://www.musee-mccord.qc.ca/fr/>.

⁵⁵ HEDGES Mark, DUNN Stuart, *Academic crowdsourcing in the Digital Humanities: Crowds, Communities and Co-production*, Londres, Chandos Publishing, Chandos Information Professional Series, 2017, p. 108-110.

amont⁵⁶. En fonction de tous ces paramètres, les outils à développer comme l'implication de l'équipe-projet au soutien au long cours de la communauté de contributeurs ne seront pas les mêmes et devront également être corrélés avec les moyens financiers alloués au projet participatif. Ainsi, pour un projet d'édition critique d'une ressource circonscrite par des experts, à l'image du projet *Digital Mellini*⁵⁷ dont l'intention était de démontrer que la collaboration de spécialistes dans des disciplines connexes à l'histoire de l'art pouvait permettre le traitement de sources difficiles d'accès (forme versifiée de registres d'inventaires d'une collection romaine constituée au XVI^e siècle), l'équipe-projet s'investira davantage dans l'accompagnement technique, la modération des débats plus que dans la vérification qualitative des contenus produits. Quand il s'agit de fonds documentaires relatifs à un sujet, un territoire ou un événement très particuliers, les communautés qui vont se former regroupent souvent des personnes qui entretiennent un lien affectif très fort avec la thématique et auront à cœur de transmettre leurs connaissances à ce sujet. Cependant, ils peuvent manquer de méthode et de rigueur scientifique pour mettre en forme leurs apports, les concepteurs du projet auront alors à charge d'assurer leur formation et de contrôler la validité des contributions pour leur assurer une reconnaissance certaine dans les milieux de la recherche académique. *Testaments de Poilus*⁵⁸ illustre ce type d'accompagnement : à l'attrait que représentent sur le public ces témoignages d'individus aux prises avec les tourments de l'histoire s'allie la nécessité pour les concepteurs de faire œuvre de pédagogie afin d'amener le contributeur à avoir la capacité de répondre aux exigences de propreté des données à générer (le choix de l'encodage en TEI-XML pour ce projet est en effet ambitieux par la complexité de l'arborescence XML et la nécessité de maîtriser un nouveau langage). Réfléchir au système de contribution à mettre en place, à la difficulté que représente la tâche pour les contributeurs et comment leur donner les moyens de répondre aux exigences qualitatives, aux étapes de post-traitement des données et à la présentation finale (édition, base de données relationnelle, etc.) demande un temps certain qui ne doit pas être sous-estimé dans le calendrier du projet.

Le volume de la collection concernée induit aussi des modes de traitement différents : une collection de 3 000 photographies ne s'indexe pas comme une collection comportant des

⁵⁶ DUN Stuart, HEDGES Mark, *Crowd-Sourcing Scoping Study. Engaging the Crowd with Humanities Research*, Centre for e-Research, Department of Digital Humanities, King's College London, Novembre 2012, p. 13, [en ligne], consulté le 15 mars 2019, URL: <https://stuardunn.files.wordpress.com/2013/04/crowdsourcing-connected-communities.pdf>.

⁵⁷ Digital Mellini : <http://www.getty.edu/research/mellini/>. Annexe I, p. 29.

⁵⁸ Testaments de Poilus : <https://testaments-de-poilus.huma-num.fr/#/>. Annexe I, p. 2.

millions d'items, la première peut être réalisée sur le site de l'institution (Archives départementales de Vendée avec le *L@boratoire des Internautes*⁵⁹) tandis que l'autre s'effectuera préférentiellement sur une interface dédiée rattachée au site du musée (*Les Herbonautes*), la commande d'un outil de traitement spécialement conçu pour la tâche (*CORRECT*, *Testaments de Poilus*) ou la réutilisation d'un service du web déjà existant ou d'un logiciel en Open Access (Flickr pour la Bibliothèque du Congrès, ScribeAPI pour *RECITAL*⁶⁰) étant fonction du budget alloué au projet. Pour Bastien Guéry, porteur du projet d'*Inventaire participatif du patrimoine national des orgues*⁶¹, « il n'existe pas de 'recette' pour la réussite d'un projet participatif en ligne, car tous les projets présentent des contraintes différentes, mais cette absence de règles à suivre n'est pas incompatible avec l'identification d'un ensemble d'écueils bien définis »⁶². Pour revenir au cas du Mobilier national, compte tenu de la diversité des publics qui s'intéressent à l'institution (curieux, passionnés d'histoire prêts à dépouiller des registres d'archives pour reconstituer l'ameublement des châteaux, chercheurs) et de la richesse des corpus, il n'est guère opportun de proposer à tous le même type d'action sur les mêmes ressources, l'idée étant plutôt de s'appuyer sur les spécificités de chacun pour créer les outils et les interfaces les plus à même de révéler le patrimoine.

1.3. *Le devenir des données produites : réintégration, réutilisation et protection*

L'enrichissement des métadonnées dans un contexte participatif n'a de sens que si le résultat est valorisé en ligne, dans des publications ou dans le catalogue des collections et permet d'améliorer l'expérience de visite du site par les internautes. Réfléchir aux conditions de visibilité du travail des contributeurs est donc crucial car ces derniers doivent percevoir l'utilité de leur participation sous peine d'être découragés et de voir faiblir leur investissement dans le projet. Cette réintégration ne doit donc pas être « cosmétique » mais réellement prévue dans le *workflow* d'alimentation du catalogue pour le cas du MN (ainsi, si un module d'indexation sociale est intégré à la plateforme MN/Lab⁶³, les mots-clés générés par les internautes doivent être intégrés au moteur de recherche du catalogue en tant que critères

⁵⁹ L@boratoire des Internautes : <http://www.laboratoire-archives.vendee.fr/>.

⁶⁰ RECITAL : <http://recital.univ-nantes.fr/#/>. Annexe I, p. 5.

⁶¹ Inventaire participatif du patrimoine national des orgues : <https://www.inventaire-des-orgues.fr/>. Annexe I, p.7

⁶² GUÉRY Bastien, «Un Inventaire participatif du patrimoine national des orgues », in *Participez / Participer. Patrimoines et dispositifs participatifs*, Journées professionnelles, Paris, 5 et 6 octobre 2017.

⁶³ MN/Lab est le nom donné au catalogue des collections du MN mis en ligne le 3 avril 2019.

permettant de créer de nouveaux chemins d'accès aux biens, plus pertinents et plus en adéquation avec les usages du public). Une veille des liens entrants (signalement d'erreurs par les contributeurs dans les notices objets) et des liens sortants (réintégration des corrections sur le site par les responsables de sa maintenance ou directement par les contributeurs) devra être accomplie⁶⁴. La réintégration des données peut se compliquer singulièrement si l'institution choisit une plateforme distincte de son propre site et préexistante : il n'y a certes pas de coûts de développement puisque tout est fourni clé en main (Flickr, ScribeAPI, Scripto) mais il n'est pas possible en revanche d'intervenir sur les fonctionnalités de rapatriement, les formats et les standards de mise en forme des données. De plus, la multiplication des plateformes peut induire un éparpillement des ressources documentaires de l'institution, à moins de disposer d'un portail d'accès bien structuré avec un instrument de recherche permettant de balayer tous les fonds (ce qui génère des coûts et des solutions de maintenance comme la création d'identifiants pérennes pour les ressources et les acteurs de la communauté). Le projet *Correct*⁶⁵ initié par la BnF illustre de tels écueils : d'une part le choix de faire développer une interface par des entreprises privées s'est révélé périlleux car celles-ci n'ont pas souhaité à l'issue des livraisons valoriser cette solution technique et l'équipe-projet s'est retrouvée avec un code impossible à maintenir seule ; d'autre part, la réintégration des corrections d'OCR par les contributeurs a été un échec car si corriger un texte page par page est en théorie aisé, la reconstruction OCR liée au PDF sur le catalogue a été impossible à effectuer et les textes corrigés n'ont pas pu être reversés sur le portail Gallica comme prévu. Quant à la fonctionnalité de recherche en texte intégral grâce aux fichiers ALTO⁶⁶, elle n'a pas abouti en raison d'une impossibilité à restituer le lien entre l'image et le texte corrigé. Des résultats mitigés pour l'institution qui a décidé à l'avenir de se tourner vers des logiciels en Open Access, quitte à faire améliorer l'outil en interne par un développeur qui aura à charge de documenter ses réalisations pour en assurer leur transmission.

Assurer la réutilisation et la protection des données générées par le *crowdsourcing* est un enjeu d'avenir pour les chercheurs sensibles à ces problématiques et certains soulèvent la

⁶⁴ MOIREZ Pauline, MOREUX Jean Philippe, JOSSE Isabelle, *État de l'art en matière de crowdsourcing dans les bibliothèques numériques*, Février 2013, Livrable L-4.3.1 du projet de R&D du FUI 12 pour la conception d'une plateforme collaborative de correction et d'enrichissement des documents numérisés, p. 76, [en ligne], consulté le 03 février 2019, URL : http://www.bnf.fr/documents/crowdsourcing_rapport.pdf.

⁶⁵ À propos de la plateforme *Correct*, consulter l'entretien avec Jean-Philippe Moreux en Annexe II, p. 53.

⁶⁶ ALTO (*Analysed Layout and Text Object*) est un standard XML permettant de rendre compte de la mise en page physique et de la structure logique d'un texte transcrit par reconnaissance optique de caractères (OCR). (Source Wikipédia).

nécessité de protocoles spécifiques qui permettraient la pérennité, l'interopérabilité et l'usage intègre de cette masse d'informations. Lisa Chupin souligne combien la génération d'édifices documentaires par des bénévoles peut entrer en conflit avec leurs ré-usages postérieurs dans des contextes interprétatifs spécialisés : quel statut juridique invoquer pour ces données générées collectivement, comment gérer les droits d'affichage et de modification, relier l'information à leur auteur si nous considérons que ce sont des individus parfaitement identifiables qui contribuent et pas une foule d'anonymes⁶⁷ ? La réponse à ces interrogations n'est pas encore advenue mais se construit à mesure que le droit s'adapte aux changements induits par le Web 2.0.

Tout d'abord, un double statut existe au regard de la loi française pour les contributions d'internautes : soit elles sont reconnues comme des œuvres originales de l'esprit et sont protégées par le droit d'auteur⁶⁸, soit elles sont considérées comme des données publiques car les activités de transcription ou d'indexation ne sont ni « originales », ni « de forme », leur auteur ne laissant pas une empreinte personnelle (les activités de co-création et d'écriture collaborative sont alors difficiles à placer dans ce paysage car le participant accepte de partager son histoire personnelle pour accroître la documentation d'un média dans une logique d'énonciation communautaire)⁶⁹. Ce sont les institutions qui vont décider des conditions générales d'utilisation des données en fonction de la nature du projet : celles édictées pour les projets *Morts pour la France de la Première Guerre mondiale* et *Correct* impliquent de renoncer à ses droits d'auteur et de céder ses droits de publication, de diffusion et de reproductions, tandis qu'une encyclopédie collaborative comme *Rosalipédie*⁷⁰ reconnaît et introduit la citation des participants. Les institutions ont aussi à charge de clarifier les conditions de protection et réutilisation des données (beaucoup de projet font le choix de demander un minimum d'informations, pseudonyme et adresse email) et de s'assurer de l'intégrité des participants qui doivent s'engager à respecter les opinions de chacun et les règles de participation. L'équipe-projet doit enfin faire preuve de mesure dans ses demandes aux bénévoles et ne pas abuser de l'enthousiasme de ceux-ci pour les surcharger de travail au

⁶⁷ CHUPIN Lisa, *Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés*, 2017, 220 p., Thèse de doctorat en Sciences de l'information et de la communication, Paris, CNAM, sous la direction de Manuel Zacklad, p. 82.

⁶⁸ Une œuvre de l'esprit doit être originale, c'est-à-dire qu'elle porte l'empreinte personnelle de son auteur, et de forme, donc énoncée selon les propres termes de l'auteur. À ce titre, elle est protégée par le Code de la propriété intellectuelle (CPI).

⁶⁹ NEROULDIS Ariane, *Le Crowdsourcing appliqué aux archives numériques : concepts, pratiques et enjeux*, Juin 2015, 109 p., Mémoire du Master Archives numériques, Lyon, Université de Lyon, sous la direction de Céline GUYON, p. 84.

⁷⁰ Rosalipédie : <https://rosalis.bibliotheque.toulouse.fr/wordpress/index.php/rosalipedie/>.

risque d'être accusée d'exploiter cette masse laborieuse à moindres coûts. La signature d'un pacte est un bon moyen de faire consensus sur les droits et les devoirs de chacun, comme *Correct* l'a fait. La charte doit porter les valeurs positives du projet et privilégier le dialogue avec l'institution, la confiance réciproque qui s'instaure étant la garantie d'un environnement de collaboration serein. Détacher spécialement un membre de l'équipe-projet pour cette modération est essentiel, elle est certes chronophage mais néanmoins nécessaire tant le lien avec les contributeurs est crucial pour la réussite du projet. C'est un conseil qui est systématiquement remonté lors de nos entretiens et que le MN s'efforce de prendre en compte dans la mise en œuvre de ses propres outils participatifs.

Sujet connexe et tout aussi sensible éthiquement : que peuvent répondre les institutions patrimoniales à l'argumentaire bien souvent avancé de participer à la dévalorisation d'un travail qui devrait être la prérogative de professionnels formés pour cela, de faire du « *dumping* salarial » et d'exploiter le sentiment d'utilité sociale qu'éprouvent les volontaires pour produire de la donnée à peu de frais ? D'un côté, il est vrai que des initiatives comme Amazon Mechanical Turk sont tout à fait répréhensibles car en plus de participer à la déshumanisation des contributeurs (aucun canal de discussion entre le patronat et les employés, manque de transparence, travail sous-payé), elles dévoient le principe même du *crowdsourcing* défini plus haut. Les musées, en demandant de générer des contenus, sont-ils à mettre sur un même plan que ces entreprises ? Mais en réfléchissant bien, ont-ils les moyens de procéder autrement pour remplir leurs missions de documentation du patrimoine, de valorisation et de diffusion ? D'un autre côté, toutes les initiatives de *crowdsourcing* à vocation économique ne se limitent pas aux pratiques de l'AMT : des formes plus ouvertes mettent l'accent sur la résolution de problèmes, le développement de solutions, l'innovation, Linux en étant une bonne illustration (tout en convenant que cela reste une forme de sous-traitance et présente un risque pour le travail salarié)⁷¹. Avec « l'économie de la contribution », nous entrerions selon Bernard Stiegler dans une troisième phase de développement du capitalisme où l'utilisateur devenu contributeur « *apporte plus que sa force de travail en mettant ses connaissances à contribution* »⁷², initiant alors de nouveaux modèles culturels et surtout de nouvelles valeurs sociétales. « *L'économie de la contribution est caractérisée par la motivation désintéressée des acteurs, orientée vers l'intérêt du plus grand nombre et la puissance de la multitude* »⁷³, ce qui rejoint finalement la théorie de « sagesse

⁷¹ HEDGES Mark, DUNN Stuart, *op. cit.* n°55, p. 106-107.

⁷² STIEGLER Bernard, in NEROULDIS Anne, *op. cit.* n°69, p. 92.

⁷³ *Idem*, p. 92.

des foules » de James Surowiecki et les valeurs d'engagement et d'altruisme énoncées par les contributeurs. Ce concept fait sens pour le patrimoine à partir du moment où les objectifs scientifiques restent au cœur du processus, que les matériaux d'études sensibles pour des questions raciales, de genre ou de sexe, ou de réutilisation au détriment de personnes et de collectifs sont accompagnés d'un dispositif pédagogique⁷⁴, et tant que les contenus créés comme les données personnelles ne font pas l'objet d'une réutilisation commerciale.

2. Des individus derrière l'écran : le rapport concepteur-contributeur, professionnel-amateur

2.1. Une communauté d'intérêt sur qui compter

Être intéressé, disponible, curieux ne suffit pas pour se prendre au jeu de la contribution, surtout lorsque les corpus à explorer sont complexes et la qualité des données à produire très fine. Un des premiers écueils à la participation des internautes est simplement le fait de connaître l'existence de ce type d'interactions avec le patrimoine, d'où la nécessité de communiquer de la part des institutions porteuses de telles initiatives. Savoir qui notre projet va attirer, comment encourager leur participation et surtout les fidéliser, est un préalable nécessaire pour déterminer la nature des activités à initier, les outils adéquats pour les réaliser, les dispositifs d'écriture pour scénariser l'ensemble, la gestion en interne des contenus et des contrôles qualité et enfin le type d'animation à créer pour que l'engouement des premiers temps ne s'éteigne pas. Pour former une communauté autour du projet, plusieurs moyens existent : les centres d'archives ont ainsi tendance à s'appuyer sur les habitués des salles de lecture pour évaluer leur intérêt et leur engagement potentiel envers la résolution de questions de recherche grâce au crowdsourcing. C'est ce que *Testaments de Poilus* a fait en sollicitant pour la recherche de ces documents dans le Minutier central des notaires de Paris et aux Archives départementales des Yvelines des généalogistes amateurs ou professionnels qui venaient régulièrement sur les sites pour leurs propres investigations. Dans le même esprit, le MN sait pouvoir compter sur un noyau dur de passionnés qui sollicitent beaucoup le département des collections pour obtenir des informations sur l'historique de certains objets, les archives afférentes à consulter et qui seraient donc concernés par une mise à disposition

⁷⁴ Resituer les corpus sensibles en spécifiant que les documents sont les témoignages d'une pensée à un moment donné de notre histoire, qu'ils sont le fruit d'un contexte politique, économique et social précis, détaché de notre actualité et donc à considérer avec recul et le plus d'objectivité possible, peut être le type de discours à tenir pour se prémunir de l'accapement d'idées à des fins de nuisance envers les communautés ou les personnes citées.

plus large des ressources. Selon l'ampleur du travail à effectuer, ce petit nombre de bénévoles peut suffire. Si ce n'est pas le cas, il est possible de faire appel à des communautés déjà constituées, telle celle active sur le réseau Zooniverse où un nombre très conséquent de bénévoles peut immédiatement se mobiliser sur les projets livrés sur la plateforme par les organismes de recherche ayant choisi la solution logicielle développée par Zooniverse et le NYPL Labs⁷⁵. Toute communauté, quelle qu'elle soit, doit être soutenue par l'équipe-projet ou avoir les moyens de s'autogérer elle-même, être tenue au courant de la progression du projet, des solutions techniques apportées en cas de problèmes, des orientations à engager pour améliorer le contenu, et donc bénéficier de canaux de communication privilégiés avec l'équipe pour que tout appel ne reste pas sans réponse. Pour *Testaments de Poilus*, ce soutien prend la forme d'un module communautaire où d'un côté les modérateurs expliquent via des commentaires pourquoi l'encodage de testament est refusé et font circuler des informations, et de l'autre, les bénévoles notifient le moment où ils quittent leur session de travail et posent des questions quand ils rencontrent des difficultés. Sans ce module, il est clairement reconnu par l'équipe que le suivi du processus serait fortement compromis, car en tablant sur une qualité très fine des rendus, les concepteurs se devaient d'offrir un accompagnement à la mesure de cette exigence. Forum, statistiques, blog, tableau des meilleurs indexeurs, portraits sont les ingrédients incontournables pour rendre visible l'effort de contribution, rassurer les bénévoles sur leur capacité à générer des contenus pertinents, mutualiser les difficultés, favoriser les apprentissages informels, promouvoir une émulation positive et exprimer la reconnaissance des concepteurs pour le travail accompli⁷⁶. Ces derniers ont d'ailleurs pour responsabilité de ne pas délaisser une communauté une fois le projet achevé : l'idéal est de pouvoir orienter les bénévoles ayant acquis des compétences techniques, analytiques, méthodologiques vers d'autres programmes afin que leur expérience puisse bénéficier à d'autres projets et que les plus férus de ce type de travail trouvent matière à nourrir ce goût⁷⁷. Mentionnons toutefois que les outils relationnels ne sont pas toujours investis comme les concepteurs l'attendaient, en ce cas, il faut s'interroger sur la manière dont ces espaces de dialogue s'articulent avec les activités proposées, l'arborescence des fonctionnalités, la hiérarchie instaurée, le contexte d'usage de la plateforme, tout ceci ayant des incidences sur le besoin des usagers de communiquer.

⁷⁵ Zooniverse : <https://www.zooniverse.org/>. Annexe II, p. 54, entretien avec Guillaume Raschia à ce sujet et description du Framework ScribeAPI dans le tableau consacré aux solutions logicielles clé en main pouvant être mobilisées dans le cadre de l'expérimentation au MN en Annexe II, p 64..

⁷⁶ CHUPIN Lisa, *op. cit.* n°67, p. 125.

⁷⁷ HEDGES Mark, DUNN Stuart, *op. cit.* n°55, p. 117.

Pour que les maîtres-mots « visibilité, popularité et interactivité » signant la réussite d'un projet de *crowdsourcing* deviennent une réalité, il faut commencer par connaître celles et ceux qui viennent spontanément vers l'institution émettrice du projet, et aussi tous ceux qui connaissent moins la structure mais qu'il faut aussi toucher, pour lesquels il va falloir réfléchir à ce qui peut les intéresser et quelles interactions créer avec eux. Une fois que nous avons une meilleure idée du profil des potentiels collaborateurs à l'amélioration documentaire de nos collections, nous devons trouver le moyen de rendre attrayantes les tâches à effectuer pour le plus grand nombre sans pour autant nuire à la qualité des données produites ni épuiser l'énergie des contributeurs dans des tâches rébarbatives qui ne mettraient pas à profit les connaissances et les compétences des bénévoles (le projet *ARTigo*⁷⁸ est parvenu à construire un juste équilibre entre attentes scientifiques d'un meilleur référencement des reproductions d'œuvres d'art par génération de mots-clés et création d'une modalité ludique de participation qui allie émulation et pédagogie autour de l'apprentissage et de la mémoire visuelle, les résultats étant plus qu'éloquents en termes d'implication et de documentarisation. Minimiser les efforts à produire (ou tout du moins faciliter les actions par la mise à disposition de toutes les aides possibles, tutoriels, modèles, ressources externes) et maximiser ainsi le temps que chacun décidera d'accorder à l'initiative ; faciliter l'accès du site à ceux qui sont les moins coutumiers des pratiques participatives (*Testaments de Poilus* s'est appuyé sur des ateliers en présentiel pour présenter le module d'encodage en XML-TEI) ; offrir une certaine flexibilité dans le choix des corpus mis en ligne (« *je souhaite décrire les planches d'herbiers de telle zone géographique parce que j'entretiens avec cette localisation un lien affectif particulier* » dans le cas des *Herbonautes* qui a aussi fait le choix d'attribuer des contenus aléatoirement aux usagers qui n'ont pas de préférence et favoriser de fait la découverte ou sérendipité⁷⁹), dans les enrichissements possibles (ajout d'images, de *tags* ou de descriptions) et dans la manière de les traiter (les contributeurs peuvent choisir de renseigner certains champs et n'ont pas pour obligation de tout remplir dans l'ordre ou lors d'une même session de travail, possibilité de travailler à plusieurs en même temps sur un même document pour s'entraider en cas de difficulté, ou inversement respect du souhait de traiter seul sur l'intégralité du document) sont aussi constitutifs du travail de préfiguration à mener⁸⁰.

⁷⁸ ARTigo : <https://www.artigo.org/about.html>. Annexe I, p. 21.

⁷⁹ Sérendipité : c'est le fait de réaliser une découverte scientifique ou une invention technique de façon inattendue à la suite d'un concours de circonstances fortuites et très souvent dans le cadre d'une recherche concernant un autre sujet (Wikipédia).

⁸⁰ SIMON Nina, *The participatory museum*, Santa Cruz, California Museum 2.0, 2010, chapitre 7, [en ligne], consulté le 04 février 2019, URL : <http://www.participatorymuseum.org/read/>.

Il n'est donc pas erroné ni anodin d'affirmer qu'à chaque projet son public et ses modalités de contributions, ses stratégies de fidélisation et ses outils. Et à chaque institution son projet participatif qui doit refléter la culture et l'esprit du musée, du centre d'archives ou de la bibliothèque, afin que le public puisse l'identifier comme lui étant particulier. Le crowdsourcing est une opportunité pour l'institution patrimoniale de s'afficher comme promotrice de créativité et d'ouverture en adhérant aux mouvements de l'Open Access et de l'Open Data qui prévalent aujourd'hui tout en conservant son identité et ses spécificités.

2.2. *La perception du crowdsourcing par les professionnels*

Rendre accessible le patrimoine conservé dans un musée, une bibliothèque ou un centre d'archives par la médiation en présentiel ou sur un site internet est une chose, permettre à l'internaute de contribuer en est une autre pouvant susciter de nombreuses craintes de la part des personnels qui y travaillent et peuvent douter de la légitimité de cette pratique, parfois perçue comme un effet de mode et non comme un processus de recherche-action à part entière. Si l'internaute peut effectivement être vecteur d'incivilités, de mésusages des plateformes, de commentaires invasifs et inopportuns, voire de manipulations de la mémoire, tous ne sont pas à ranger dans cette catégorie et peuvent au contraire apporter une réelle plus-value aux contenus numériques et collaborer en bonne intelligence avec les professionnels sans remettre leur statut et leur expertise en cause⁸¹. Pour Serge Noiret, les principes interactifs du Web 2.0 sont faits de pratiques qui « *défient les lois classiques du métier [d'historien], sans remettre en cause son ontologie même* » et incitent « *à réviser les comportements épistémologiques de nombreuses disciplines humanistes* »⁸². Le professionnel pourrait alors apparaître comme le vecteur le plus à même de favoriser la rencontre entre le patrimoine et le public, tout en étant le garant de la qualité scientifique des contenus produits, sa formation, son regard et sa maîtrise des outils de recherche et des normes lui permettant d'instaurer une éthique de collaboration qui respecte les individus et leurs acquis. Mais cette posture bute contre l'idée que le travail des documentalistes et des conservateurs se trouve déprécié par cette contribution gratuite, effectuée sur le temps du loisir, émanant de toutes parts sans parfois avoir la possibilité d'identifier les individus derrière le commentaire. Le sentiment de perdre le contrôle sur les contenus générés s'accroît encore avec la numérisation

⁸¹ MOIREZ Pauline, MOREUX Jean Philippe, JOSSE Isabelle, *op. cit.* n°64, p. 74.

⁸² NOIRET Serge, « La Digital History: histoire et mémoire à la portée de tous », in *Ricerche Storiche*, 1, vol. 41, Janvier-Avril, 2011, p. 138.

à la demande où ce ne sont plus les bibliothécaires qui ont la main sur les corpus mis en ligne mais les usagers⁸³. De même, mettre en ligne les collections d'un musée pose la question de l'usage des images et des informations à leur sujet (le choix de libérer les droits sur les images n'est pas anodin, c'est un parti pris que le MN a fait mais que tous les musées ne sont pas enclins à assumer). Un de nos interlocuteurs s'est élevé contre la pratique visant à transcrire des fonds d'archives pour en faciliter l'usage par un plus large nombre de chercheurs et d'amateurs, arguant que dépouiller des cartons, déchiffrer des écritures anciennes, recontextualiser est constitutif du travail de recherche, c'est un apprentissage, une « expérience de terrain » qui ne doit pas être supprimée sous prétexte d'outils et de modes de diffusion nouveaux. Un travail sur des numéros d'inventaire qui, aux dires de Pascal Bertrand, Professeur d'histoire de l'art à l'Université de Bordeaux-Montaigne, serait très précieux au sujet des collections du MN pour mieux connaître les logiques d'attributions de ces marques⁸⁴, est, pour Yves Carlier, conservateur général du département des ressources documentaires du château de Versailles, un outillage et une production qui doivent rester en interne et n'ont pas pour vocation à être diffusés et effectués par des internautes.

Ce qui distingue l'amateur du professionnel, c'est la maîtrise par ce dernier des normes de catalogage et de description bibliographique, des règles d'indexation (constitution de thesaurus, listes d'autorité, classifications) et de transcription diplomatique, des standards d'encodage (EAD, TEI) et des modèles de descriptions données (HADOC, CIDOC-CRM). Donner les clés de compréhension et d'usage de ces savoirs à tous est perçu par certains archivistes et documentalistes comme une dévalorisation du temps et des efforts qu'ils ont alloués à l'acquisition de ces compétences lors de leurs études, voire comme une mise en danger de tout un corps de métiers. C'est le monde scientifique qui se trouve questionné par ce qu'Antonio Casilli appelle le *Digital Labor*, et la « *pression salariale [qui] s'exerce dans les secteurs les plus divers (journalisme, industries culturelles, transports, etc.) par la création d'une armée de réserve de 'travailleurs qui s'ignorent', convaincus d'être plutôt des consommateurs, voire des bénéficiaires de services gratuits en ligne* »⁸⁵. Pour apaiser les tensions que suscite le *crowdsourcing*, Mathieu Andro appelle à une médiation où l'expertise

⁸³ ANDRO Mathieu, *Bibliothèques numériques et crowdsourcing : Expérimentations autour de Numalire, projet de numérisation à la demande par crowdfunding*, 2016, 461 p., Thèse de doctorat en sciences de l'information et de la communication, Saint-Denis, Université Paris 8 Vincennes, sous la direction de Imad Saleh.

⁸⁴ Nous reviendrons sur ce point plus en détail dans la Partie III.

⁸⁵ CASILLI Antonio A., CARDON Dominique, *Qu'est-ce que le Digital Labor ?*, INA éditions, 2015, in HERBILLE Gauthier, MAZET Jérémy, PETIT Axel, « Transcribe Bentham : recherche historique et crowdsourcing », in *Innovation pédagogique*, 3 janvier 2017, p. 16, [en ligne], consulté le 3 mars 2019, URL : <https://www.innovation-pedagogique.fr/article1329.html>.

des professionnels des métiers de la documentation serait réaffirmée auprès des internautes et où l'ouverture vers d'autres points de vue et parcours serait mise en exergue pour que les professionnels restent en phase avec les aspirations de la société et des individus qu'ils sont censés servir par la mise à disposition d'une information stable, pérenne et interopérable⁸⁶. Enfin, laisser la possibilité aux équipes de l'institution d'exprimer leurs craintes et leurs attentes pour que le projet soit un gain pour elles, une manière de valoriser leur travail et non une charge jugée vaine et contre-productive scientifiquement parlant, chronophage et peu porteuse pour l'image de l'institution, sont également un bon moyen de désamorcer conflits et incompréhensions.

3. Des données, des compétences, des outils : comment concevoir le modèle de *crowdsourcing* adapté ?

3.1. Les étapes de la conduite de projet

S'il n'existe pas de protocole de création d'un projet de *crowdsourcing* type transposable d'un contexte à un autre, il est néanmoins possible d'émettre des recommandations après le recueil d'expériences et l'analyse de quelques initiatives. Il convient tout d'abord pour les porteurs du projet d'affirmer auprès de leurs collègues que cette modalité de mise en valeur des collections s'inscrit en conformité avec le schéma directeur de l'institution et n'est pas un coup de publicité sans fondements scientifiques derrière. Dessiner des objectifs précis, circonscrits, pouvant être atteints à courte ou moyenne échéance, communiquer sur les apports de la démarche et sur les réalisations effectives sont de bons moyens pour démontrer la pertinence de l'initiative, la réflexion qui est à sa base et finalement désamorcer les réticences éprouvées. En demandant à chacun ce que pourrait apporter une meilleure documentation des collections pour son travail quotidien, nous pouvons définir leurs besoins et envisager l'activité la plus à même d'y répondre⁸⁷. Et parallèlement, en accordant à chacun la liberté d'exprimer son opinion et son intérêt pour le sujet, nous pouvons aussi faire émerger des volontés qui seraient prêtes à aller au-delà de ce rôle consultatif pour s'engager plus avant dans le projet, y consacrer du temps et finalement rejoindre l'équipe qui va devoir se constituer et être clairement identifiable pour que le projet prenne corps dans la politique

⁸⁶ ANDRO Mathieu, *op. cit.* n° 83, p. 272.

⁸⁷ La structuration des données demande une coordination et une certaine dose d'agency de la part des professionnels qui doivent avoir à l'esprit qu'une première tâche proposée aux participants donnera une première couche de sens au corpus qui pourra donner lieu à une seconde tâche plus complexe en termes de description des données et d'outils pour ce faire, et ainsi de suite à mesure que les objectifs évoluent, se précisent dans une direction ou une autre.

générale de l'établissement. En fonction de ces engagements, les ressources budgétaires, les moyens humains (évaluation du temps de travail que représentera la conduite du projet sur les autres missions que doit remplir un agent, recrutement éventuel) seront chiffrés, la recherche de partenariats ou la réponse à des appels à projet (comme ceux émis par l'Agence Nationale de Recherche ou les Entrepreneur.se.s d'Intérêt Général) s'avérant indispensables pour se lancer car les institutions patrimoniales souffrent bien souvent d'un manque de personnel pour atteindre leurs objectifs premiers et ne possèdent pas forcément les compétences techniques pour développer des outils informatiques. Concomitamment à cela, un tour d'horizon des projets en ligne (*benchmarking*) qui présentent des caractéristiques communes au nôtre en termes de corpus d'étude, de tâches⁸⁸, de partis pris méthodologiques, de procédures de modérations des contributions et de publics touchés peut aider à la préfiguration du projet, donner un aperçu des pratiques qui fonctionnent et des écueils auxquels penser pour trouver des solutions. Cibler les médias les plus adéquats pour assurer la promotion et le rayonnement du programme s'exécute également dans cette temporalité. Enfin, établir dès cette phase de préfiguration les modalités d'évaluation du projet est important pour avoir les moyens de réagir en cas de stagnation de la participation après la phase de lancement.

Une fois ces jalons posés, il est possible de passer à la phase de conception proprement dite de la plateforme, à savoir son architecture technique, fonctionnelle et ergonomique. Selon les objectifs visés et les moyens octroyés, l'équipe choisit de créer de toute pièce un assemblage logiciel parfaitement adapté ou d'utiliser une formule existante et déjà éprouvée. À chaque étape de développement, à chaque addition de fonctionnalité, il est plus que recommandé (et souvent très bien accueilli) de solliciter le personnel de l'institution et un petit panel d'internautes⁸⁹ pour tester les versions, faire remonter les remarques et corriger au fur et à mesure le prototype (ce qui correspond à l'application des méthodes Agiles). De surcroît, toute modification de l'outil, les procédures et les versions de la plateforme doivent absolument être documentées par les développeurs car si jamais l'entreprise en charge de la création logicielle connaît des avatars mettant en péril la poursuite de ses activités, et donc la maintenance de son service, il doit être possible de reprendre le travail là où il a été laissé (une clause dans le cahier des charges peut obliger la spécification des conditions de ce transfert de compétences par l'entreprise qui remportera l'appel d'offre) et ne pas mettre en péril le projet

⁸⁹ Le MN a proposé aux personnes sollicitant le plus souvent le personnel de la documentation de devenir les testeurs de la version bêta du catalogue des collections, la pertinence des retours ayant permis de faire évoluer le modèle pour répondre au mieux aux attentes du plus grand nombre.

dans son ensemble⁹⁰. Une solution très simple consiste à recourir au service d'hébergement et de gestion de développement de logiciels *GitHub*⁹¹ pour signaler les améliorations à prévoir (tickets), celles en cours de traitement et celles qui ont été réalisées à tous ceux et celles qui suivent le projet, la description de chaque brique logicielle étant également consultable pour garantir une totale transparence, la réutilisation et les évolutions futures de l'outil.

En phase industrielle, c'est-à-dire durant le fonctionnement effectif de l'interface, l'effort de l'équipe-projet sera de maintenir la participation des contributeurs recrutés dans les premières heures du projet. Il est possible d'agir sur plusieurs plans pour cela : évolution des tâches et des corpus sur lesquels travailler en adéquation avec les demandes du public et les besoins de l'institution, réactivité de l'équipe à répondre aux interrogations des internautes et encouragement des contributeurs à prendre en main le projet par l'attribution de statuts particuliers (vérificateur, chef de mission)⁹². La capacité à écouter les bénévoles a ainsi été relevée (et appréciée) par Fabien Laure, indexeur aguerri de la base des *Morts pour la France de la Première Guerre mondiale*, qui s'intéressait particulièrement aux Poilus vosgiens et souhaitait tous les indexer. Constatant l'absence de visuels de fiches pour achever le traitement des combattants de cette région, il sollicita le webmestre de la plateforme qui a pu débloquent rapidement et spécialement pour lui la quarantaine de fichiers manquants. Pour indication, Fabien Laure était classé au rang 22 sur 2331 en janvier 2018. En relayant la demande de ce contributeur, le webmestre illustre l'intérêt d'encourager les volontaires les plus assidus par des actes concrets (mise à disposition de documents, réponses rapides) car ce sont eux qui feront le plus progresser le projet par leur capacité de travail colossale⁹³. Ce qui nous amène à la dernière phase d'un projet, la phase d'analyse. Elle comprend la réintégration des données produites (plan de migration des données, voir Partie I), leur mise à disposition en suivant la politique d'ouverture des données intrinsèquement liée à ce type de projet (l'indexation des contenus par Google étant le choix le plus souvent retenu) et l'évaluation du projet, voire sa critique. Travailler en « agilité » est donc plus que prescrit et s'inscrire dans « *une démarche de recherche-action qui permet d'appréhender le développement de pratiques émergentes et les mutations des formes de communication médiatisées par les dispositifs documentaires en cours de transformation* »⁹⁴ largement plébiscité.

⁹⁰ Voir l'entretien avec Karl Pineau le 9 mai 2019, Annexe II, p. 56.

⁹¹ Les EIG engagés par le MN ont créé une session pour la mise en ligne des collections consultable à cette adresse : <https://github.com/entrepreneur-interet-general/gobelins>.

⁹² MOIREZ Pauline, MOREUX Jean Philippe, JOSSE Isabelle, *op. cit.* n°64, p. 72.

⁹³ Voir le portrait de Fabien Larue sur le portail *Mémoires des Hommes* du Ministère des Armées.

⁹⁴ CHUPIN Lisa, *op. cit.* n°67, p. 55-56.

3.2. *S'assurer de la qualité des données produites : la question charnière de la validité scientifique du projet*

En fonction de la nature de l'activité proposée, de l'exigence de qualité attendue et de la confiance accordée aux contributeurs par les institutions émettrices, les procédures de validation des contenus générés dans le cadre d'un projet de *crowdsourcing* sont très diverses. De la prise en main complète de l'outil par les internautes qui modèrent les contributions des uns et des autres au contrôle strict des données produites par des experts affiliés au projet, l'équipe doit choisir le procédé, tout en sachant que « *plus l'encadrement [des contributions] est fort en amont, moins il l'est en aval, et réciproquement* »⁹⁵. Ceci signifie par exemple que même si l'enthousiasme qu'a soulevé le projet d'indexation collaborative *Morts pour la France de la Première Guerre mondiale* a donné un nombre très conséquent de fiches à vérifier au webmestre (350 000 notices à raison d'une dizaine de champs par notice à renseigner), en favorisant les menus déroulants où il faut cocher l'intitulé correct (donc la forme attendue par l'équipe) et les champs semi-libres à auto-complétion (l'internaute saisit les premiers caractères et l'interface suggère une expression en s'appuyant sur les thesaurus préenregistrés), le travail de vérification est grandement facilité par alignement automatique des réponses. Tout en étant chronophage, les concepteurs ont trouvé dans cette procédure un palier de sécurité fort appréciable et un rempart contre le risque d'erreurs, écueil bien souvent souligné dans les critiques émises à l'encontre du *crowdsourcing*. Les contributeurs doivent aussi se créer un compte, ce qui peut constituer un premier obstacle dissuasif à tout acte de malveillance, et s'engager au respect de la charte du projet qui figure en page d'accueil⁹⁶.

La formation des contributeurs est une autre manière d'assurer une production qualitative, les outils pédagogiques développés par *Testaments de Poilus* sont à ce titre exemplaires : grâce à des tutoriels explicitant chaque étape du processus et montrant le détail de ce que permet chaque balise et à quel moment les utiliser, des novices ont appris à coder en XML-TEI et ont réalisé un travail d'une grande finesse (à tel point que pour les meilleurs codeurs, un rôle de vérificateur a été créé pour mettre à profit leurs compétences). De même, *les Herbonautes* ont mis à disposition des bénévoles un lexique afin de les aider à normaliser les données saisies et optimiser leur participation, des index répertoriant des informations sur les

⁹⁵ ZACKLAD Manuel, CHUPIN Lisa, « Le *crowdsourcing* scientifique et patrimonial à la croisée de modèles de coordination et de coopération hétérogènes: le cas des herbiers numérisés », in *Canadian Review of Information and Library Science*, Toronto, University of Toronto Press, 39 (3/4), 2015, p. 321, [en ligne], consulté le 04 avril 2019, URL : <https://muse.jhu.edu/login?auth=0>.

⁹⁶ Engagement de l'annotateur en sept points pour le projet *Morts pour la France de la Première Guerre mondiale* : <https://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?larub=52&titre=annotation-collaborative>.

botanistes, le vocabulaire descriptif des plantes et des liens vers d'autres bases de données (Sonnerat/BryMyco et JSTOR Global Plant) dont les herbiers numérisés peuvent aider à identifier les spécimens. Ce projet recourt aussi à des tests de niveaux pour évaluer les progrès des contributeurs et les autoriser à renseigner de plus en plus de champs (et donc à endosser des responsabilités plus importantes au sein du programme), marque de la reconnaissance des concepteurs de l'expertise et de la capacité des usagers à dialoguer et résoudre les cas litigieux ensemble. Le classement des documents selon leur difficulté de traitement et leur allocation aux « meilleurs » contributeurs est aussi une stratégie pour avoir la présomption, sinon la certitude, que la source sera traitée de manière la plus optimale possible par la bonne personne. L'idée d'attribuer les corpus aux individus les plus à même de les exploiter est une fonctionnalité sur laquelle des chercheurs en ingénierie informatique et en sciences humaines travaillent pour améliorer l'acte participatif.

Lorsque les moyens octroyés au projet sont insuffisants pour désigner une personne en charge de la modération et de la validation des contenus de la plateforme, et si les activités comme les corpus le permettent, la validation collaborative s'impose. Elle sous-tend la confiance de l'équipe envers la communauté pour accepter de déléguer la vérification, mais cette vérification peut aussi être implicite, les contributeurs corrigeant sans le savoir des données, excluant de fait les risques de tricherie⁹⁷. Cette modalité, nous la rencontrons souvent en dehors de tout contexte collaboratif lorsque nous nous connectons sur notre adresse email : c'est le ReCaptcha. Lorsque le projet Google Books a commencé à numériser en masse des livres pour en permettre la libre consultation en ligne, nous avons constaté un OCR de piètre qualité (25 % des mots océrisés comportent des erreurs) qu'il fallait pouvoir améliorer aisément. Pour cela, les mots sont envoyés sur le système ReCaptcha qui va proposer à trois internautes lors de leur connexion sur un compte le même mot avec trois déformations différentes. Si les trois saisies des internautes sont identiques, la correction du mot est validée et ainsi de suite. En cas de corrections différentes, le mot est proposé à d'autres internautes jusqu'à obtenir un taux de vote de 2,5 pour entériner la correction⁹⁸. Utiliser un outil conçu pour la lutte contre les robots malveillants afin de rendre accessible des documents patrimoniaux montre que les sciences humaines et l'informatique ne sont pas des disciplines antinomiques mais peuvent au contraire dialoguer pour perfectionner les outils du

⁹⁷ PINEAU Karl, « CaptchA-N : Vers un système de validation de l'indexation collaborative », in *Données culturelles. Carnet de recherche sur les données et ontologies patrimoniales*, publié le 11 décembre 2018, [en ligne], consulté le 14 février 2019, URL : <https://dculturelles.hypotheses.org/33>.

⁹⁸ ANDRO Mathieu, *op. cit.* n° 83, p. 157.

chercheur. Toutefois, il n'est pas recommandé de laisser le contributeur en dehors du processus de validation, une certaine transparence des corrections effectuées doit demeurer afin de ne pas ériger de barrière entre les concepteurs et les contributeurs, la vocation des sites de *crowdsourcing* en sciences humaines étant de créer du lien entre dépositaires du savoir. La conservation de l'historique des modifications et la restauration des propositions antérieures permettent enfin de garder un œil sur la qualité des données produites et sur les améliorations à prévoir en termes d'autorégulation. TROVE et Wikisource fonctionnent selon ce procédé⁹⁹.

En favorisant le vote, le croisement des regards et des points de vue, certaines plateformes ont choisi de miser sur la sagesse collective et sur l'engagement des contributeurs envers le projet à assurer la validation des données produites. Cette solution peut favoriser l'émergence d'une véritable appropriation de la démarche de recherche scientifique par les internautes qui vont devoir faire preuve de rigueur pour justifier leurs réponses, développer des argumentaires pour défendre leur posture en cas de dissensus, donc ne pas se contenter de ce qui semble juste mais l'est assurément (jusqu'à preuve du contraire) par l'analyse et la prise de recul. Le projet *Biodiversity Heritage Library*¹⁰⁰ valide ce postulat : la tâche, mécanique en soi puisqu'il s'agit de tagger des illustrations sans obligation formelle d'expliquer l'identification proposée, demande néanmoins au contributeur de citer les sources l'amenant à penser que telle caractéristique morphologique chez cet animal ou ce végétal indique son appartenance à telle espèce. C'est au contributeur de montrer sa fiabilité aux autres participants qui pourront, après un temps d'évaluation, considérer ses apports comme signifiants scientifiquement parlant. Une procédure externe de contrôle n'est alors pas nécessaire à créer, la communauté exerçant spontanément cette action.

Pour que cette autonomie de la communauté soit acceptable aux yeux des professionnels, certains ont entériné le fait de découpler la base de données originale à l'origine du projet de celle qui va être enrichie par les annotations des bénévoles : les professionnels ont la possibilité de valider les contenus s'ils le souhaitent et décident ensuite d'importer dans la base institutionnelle les corrections quand ils le jugent bon. C'est une structuration de l'information plus souple qui permet de « *préserver l'autorité institutionnelle sur les collections sans compromettre la dynamique de participation* »¹⁰¹. C'est la raison pour laquelle la Bibliothèque du Congrès et de nombreux centres d'archives français ont choisi

⁹⁹ MOIREZ Pauline, « Bibliothèques, catalogues et crowdsourcing », in *Archives masala*, 20.10.2014 (extrait du cours sur Les évolutions des catalogues, ENSSIB, 6.10.2014), [en ligne], consulté le 09 janvier 2019, URL : <https://archivesmasala.wordpress.com/2014/10/20/bibliotheques-catalogues-et-crowdsourcing/>.

¹⁰⁰ Biodiversity Heritage Library: <https://www.biodiversitylibrary.org/>. Annexe I, p. 40.

¹⁰¹ CASEMAJOR LOUSTAU Nathalie, in CHUPIN Lisa, *op. cit.* n°67, p. 111.

Flickr pour permettre aux internautes de construire et partager des connaissances à propos de leurs collections car ces apports s'effectuent sur une interface clairement distincte du site institutionnel, avec ses propres règles de participation et de validation, l'institution gardant en dernier ressort la main sur la réintégration des contenus et sur la modération des discussions.

Finalement, que nous options pour une assurance-qualité fondée sur la rétroaction (par des professionnels ou des participants, automatisée ou manuelle), sur la tâche (formation, sélection des volontaires) ou sur le produit¹⁰² (comparaison des contributions, agrégation des résultats, réponses multiples), il s'agit bien de penser ce choix en fonction des particularités de notre projet et si nous faisons parfois des compromis, aucune compromission ne doit en revanche être tolérée, la reconnaissance scientifique du projet étant en ligne de mire.

3.3. *Estimer la réussite d'un projet et sa viabilité : sur quels critères baser cette analyse ?*

Nathalie Casemajor Loustau souligne la difficulté à établir des critères d'évaluation autres que quantitatifs (nombre de notices produites, de lignes transcrites) des projets participatifs, surtout quand le projet repose sur le partage d'histoires personnelles dont juger la véracité est très délicat¹⁰³. Pourtant, mesurer l'utilité du travail réalisé par les contributeurs et les apports réels pour la documentarisation est important pour que le *crowdsourcing* ne soit pas seulement une modalité de communication, une expérimentation d'interaction avec le public sans devenir. Le recours au service gratuit d'analyse d'audience d'un site web *Google Analytics* qui permet notamment le suivi d'événements ou la création de tableaux de bord, la réalisation d'enquêtes comme le fit l'équipe du projet *Correct* pour connaître les contributeurs et leur utilisation des outils de communication à leur disposition, ou encore les entretiens en direct avec les contributeurs (Portraits du site *Morts pour la France de la Première Guerre mondiale*) sont quelques solutions éprouvées pour obtenir quelques données à exploiter mais elles sont insuffisantes pour identifier clairement les facteurs de réussite ou d'échec d'un projet. Mia Ridge souligne ainsi la nécessité pour l'équipe-projet d'avoir à l'esprit dès la phase de conception les objectifs à atteindre, ce que cela va changer en termes de documentation et de valorisation des collections, d'inclusion du public dans la politique de

¹⁰² VIKHROVA Anne, *L'évaluation de la méthode du crowdsourcing pour la transcription de manuscrits*, 2017, 331 p., Thèse, Laboratoire de linguistique et didactique des langues étrangères et maternelles, sous la direction de Thomas Lebarbe, p. 289-291.

¹⁰³ CASEMAJOR LOUSTAU, Nathalie, « La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? », in *Études de communication*, 36, 2011, p. 6, [en ligne], consulté le 31 mars 2019, URL : <http://journals.openedition.org/edc/2532>.

l'établissement, de mobilisation du personnel autour d'initiatives qui impliquent une certaine flexibilité dans la manière de concevoir leur travail dans un monde profondément modifié par les technologies numériques et le Web 2.0¹⁰⁴. Établir de tels indicateurs est une opération ardue car elle demande de réelles capacités d'abstraction pour imaginer comment évaluer des résultats encore non avenus et incertains du fait que ce sont des êtres humains qui participent et pas un automate dont les actions sont paramétrables. D'où l'importance des phases de tests pour observer la réception de l'information par les professionnels et les amateurs et l'usage qu'ils en font et pour mesurer l'impact réel des facteurs de complexité repérés en amont. Que les difficultés apparaissent au niveau du contributeur (expérience de la transcription, de l'encodage en XML-TEI, des outils numériques, formation professionnelle, profils socioculturels, temps alloué, etc.), du schéma descriptif (usage de vocabulaire contrôlé ou libre, degré de précision des descriptions, hiérarchisation des informations) ou du document lui-même (nombre de lignes, de caractères spéciaux, écriture plus ou moins lisible), elles sont à prendre en compte dans l'appréciation de l'avancement du projet¹⁰⁵.

Enfin, et il s'agit certainement là d'une manière d'achever cette deuxième partie consacrée à la mise en œuvre concrète d'un projet de *crowdsourcing* au sein de GLAMs, il convient de s'interroger sur la viabilité de ce type de projets d'humanités numériques, l'intérêt pour les institutions patrimoniales de les maintenir sur le long terme en sachant l'investissement humain et financier que cela demande ? Olivier Le Deuff insiste sur le devoir d'instaurer une distance critique avec ce mode de construction du savoir : diffuser l'information ne signifie pas nécessairement construire de la connaissance à son sujet, les risques de récupération idéologique et économique étant toujours présents¹⁰⁶. Inciter les concepteurs comme les contributeurs à construire des parcours de transmissions du savoir plutôt que des schémas strictement ordonnés et préétablis donnera une marge de manœuvre pour faire évoluer le projet au fil du temps, y intégrer d'autres visions et d'autres modalités de fonctionnement qui assureront la continuité de l'initiative.

¹⁰⁴ RIDGE Mia, « Playing with Difficult Objects – Game Designs to Improve Museum Collections », in J. Trant and D. Bearman (eds), *Museums and the Web 2011: Proceedings*, Toronto: Archives & Museum Informatics, 31 mars 2011, p. 72, [en ligne], consulté le 19 février 2019, URL : http://conference.archimuse.com/mw2011/papers/playing_with_difficult_objects_game_designs_improve_museum_collections.

¹⁰⁵ VIKHROVA Anne, *op. cit.* n°102, p. 294-295.

¹⁰⁶ LE DEUFF Olivier, in LEBRETON Anne, *Bibliothèques, tags et folksonomies. L'indexation des bibliothèques à l'ère sociale*, mars 2008, 117 p., Mémoire d'étude, Paris, Département des ressources documentaires de l'ENSSIB, sous la direction de Nicolas MORIN, p. 75.

TROISIÈME PARTIE : L'EXEMPLE DU MOBILIER NATIONAL

Le participatif en gestation

1. S'ouvrir au plus grand nombre : un défi à la mesure d'une grande institution patrimoniale

1.1. L'Entreprenariat d'Intérêt Général (EIG) : un tremplin

À la suite de riches échanges avec les étudiants de l'École nationale supérieure de création industrielle (ENSCI), des quatre écoles d'art appliqué de Paris et avec le Secrétariat général pour la modernisation de l'action publique (SGMAP), et en conformité avec les souhaits émis par celles et ceux qui travaillent au sein de l'institution, la nécessité de multiplier les accès au patrimoine matériel et immatériel du MN est devenue prégnante, cette ouverture pouvant être à la fois physique et numérique. Dès 2014, une note de cadrage promeut la création d'un portail en ligne des collections ; incluse dans le schéma directeur de l'établissement, elle est ensuite reprise et fusionnée avec le projet des musées que conduit alors le Service des Musées de France (SMF). Malgré tout l'intérêt de la démarche, le volet de mise en ligne des collections est écarté en 2016 du projet car il est jugé secondaire dans l'ordre des priorités édictées par le SMF. Le MN décide alors de passer par d'autres canaux pour rechercher les compétences et les financements, la réponse aux appels à projet étant une piste privilégiée. Après un long travail préparatoire pour connaître l'existant en termes de mise en ligne de collections et de ressources et les besoins des professionnels de l'institution et d'ailleurs, Jérôme Poulain, secrétaire général du MN et mentor stratégique, et Hélène Cavalié, directrice adjointe des collections et mentor opérationnel, ont été auditionnés en 2017 au Secrétariat d'État à l'économie numérique pour promouvoir le projet d'ouverture des données du MN dans le cadre du programme EIG. Lauréats de cet appel à projet, le MN a pu recruter un développeur, Ned Baldessin, et une designer, Laurie Chapotte, pour élaborer la future plateforme numérique. Les atouts d'un développement en interne de ce type d'outils sont considérables : outre l'engagement de ces deux experts à proposer une solution conforme aux

objectifs définis en amont par le MN dans un laps de temps contraint (dix mois), gage d'une émulation et d'une réactivité constantes, le fait d'être présents physiquement au MN, d'assister aux réunions des différents services, a permis de prioriser les fonctionnalités qui sont essentielles pour les agents qui ont partagé leur vision du travail qu'ils effectuent au MN, de comprendre les particularités des missions que poursuit cette institution sans égal dans le paysage muséal français et de proposer les solutions techniques, ergonomiques et éditoriales les plus adéquates aux valeurs que tous souhaitent véhiculer. Si nous n'allons pas entrer plus dans le détail de cette mise en ligne des collections, nous ajouterons toutefois que tout le travail de préfiguration réalisé par les équipes du MN s'est trouvé augmenté par le recours à la méthode de conduite de projets Agile qui a prouvé dans ce contexte toute sa pertinence.

À la différence d'une gestion de projet classique où les chefs de projet établissent un planning contraint et non évolutif, les livraisons des rendus étant toutes prévues en amont sans aucune souplesse, la méthode Agile propose de segmenter les fonctionnalités essentielles à développer (ici, une interface de recherche, une liste de résultats, une notice détaillée) sous la forme de mini-projets avec leurs objectifs propres, leurs budgets et leurs tests, pour qu'en cas de changement d'orientation, tout le projet ne soit pas remis en question mais seulement une ou deux briques qui demandent à être repensées. Cette flexibilité est rendue possible par la concertation permanente du client avec l'équipe-projet, celle-ci étant très attentive au vécu et aux besoins que les professionnels décrivent, des phases de tests régulières venant confirmer ou infirmer la pertinence de la solution proposée. Concrètement, les EIG imaginent une réponse au problème posé, font une première livraison qui va être discutée par les personnes concernées, les EIG réagissent, des itérations sont effectuées régulièrement et évaluées par les premiers concernés par la question. Une fois que la fonctionnalité souhaitée est considérée comme atteinte collégialement, une date est fixée pour la livraison de la fonctionnalité suivante à déployer. L'idée de cette démarche est de mêler étroitement l'expertise de l'utilisateur qui connaît parfaitement le contexte de l'institution et sera à même de mesurer l'adéquation de l'outil avec celle-ci, et de ne plus avoir à patienter jusqu'au terme du projet pour l'évaluer. En effet, quand l'utilisateur évalue l'outil en fin de développement, si jamais il s'aperçoit que des fonctionnalités ne conviennent pas à sa pratique professionnelle, toute modification va entraîner de tels surcoûts que ses remarques risquent de ne pas être prises en compte et finalement l'outil ne correspondra pas à ce qu'il avait espéré pour améliorer son travail. Avec la méthode Agile, le fait d'évaluer chaque fonctionnalité l'une après l'autre, de corriger immédiatement ce qui ne va pas et seulement à cette condition de passer à l'étape

suiuante permet de rester dans les limites budgétaires prévues au début, de garder une cohérence dans la gestion du projet et de ne pas mettre en péril celui-ci par un changement de prestataire à l'issue de la livraison finale. Ces caractéristiques font de la création du catalogue des collections du MN un modèle en son genre, qui mérite d'être exporté et appliqué dans d'autres institutions culturelles pour son efficacité et sa logique collaborative fédératrice.

1.2. Concerner le plus grand nombre : comment et pourquoi impliquer le public dans la vie de l'institution ?

Si le nom du Mobilier national est connu, en revanche ses missions, elles, le sont moins, ou tout du moins, seule une partie est transparente aux yeux du grand public. Pour remédier à ce constat, une véritable politique de communication et d'ouverture des lieux, aussi bien physiquement que digitalement, a été entreprise depuis deux ans. Outre les expositions temporaires dans la galerie des Gobelins, l'accueil de festivals, les Journées Européennes des Métiers d'Art, les Rencontres des Gobelins et les visites des manufactures¹⁰⁷, le MN a ouvert physiquement les portes de ses réserves et des ateliers pour la première fois en septembre 2018 lors des Journées européennes du Patrimoine. Le public s'est déplacé en grand nombre pour découvrir l'institution, les coulisses de celles et ceux qui œuvrent au quotidien pour l'ameublement des lieux d'exercice du pouvoir et la transmission des savoirs et savoir-faire permettant de conserver un patrimoine d'exception et de créer. Pour renforcer cette opération et insister sur le volet « diffusion », le 21 septembre 2018 les ateliers de restauration du MN ont accueilli des Wikimédiens venus écouter les lissiers, ébénistes, menuisiers en siège, lustriers et bronziers parler de leur métier, montrer leurs outils, les gestes les plus courants... Des photographies et des vidéos ont été collectées et diffusées sur la plateforme de l'association Wikimedia Commons¹⁰⁸ pour être librement partagées et réutilisées. L'objectif de cette médiation était de rendre perceptible le travail de ces artisans et surtout de montrer que ces savoirs ne sont pas conservés jalousement entre les murs de l'institution mais ont vocation à être diffusés et transmis.

Lorsque le projet de mise en ligne des collections est devenu de l'ordre du possible, il était nécessaire de déterminer quels publics deviendraient les utilisateurs de l'interface pour que l'offre soit en correspondance avec leurs attentes. Pour cela, des Personas ont été créées, c'est-à-dire des portraits-type de personnes dont le métier, les loisirs, l'attachement au

¹⁰⁷Cultural, Ateliers des Gobelins : <https://www.cultural.fr/visites/ateliers-des-gobelins-dans-les-coulisses-dun-metier-dart>.

¹⁰⁸ Photos et vidéos : https://commons.wikimedia.org/wiki/Category:Atelier_Mobilier_national_du_21_septembre_2018.

patrimoine pouvaient être des raisons les amenant à consulter le catalogue. Ces Personas sont inspirées d'individus qui travaillent au MN ou qui ont contacté à plusieurs reprises les équipes pour des demandes particulières. Ainsi, nous trouvons un historien d'art chercheur au CNRS qui veut consulter les archives, un professionnel du MN en charge de la mission d'ameublement qui doit faire des propositions en fonction des critères très précis émis par le chef de cabinet d'un ministre, une designer d'espace qui cherche des idées et veut télécharger des images en HD pour créer des modèles, un restaurateur qui veut accéder à des informations sur la confection de l'objet qu'il doit traiter, une commissaire d'exposition qui souhaite trouver des ressources et pourquoi pas des objets à emprunter pour illustrer son propos, une passionnée de loisirs textiles qui apprécierait de trouver des modèles pour ses confections, une étudiante en histoire de l'art qui cherche à mieux distinguer les styles, les ornements caractéristiques, et bien d'autres encore¹⁰⁹. Pour chacune de ces Personas, l'aisance numérique, la connaissance du MN et de ses ressources, et la fréquence de l'usage de l'interface ont été estimées (note de 1 à 5) afin d'évaluer la facilité à utiliser l'informatique pour trouver des éléments de réponse, l'ergonomie la plus agréable qui doit être trouvée pour la majorité et les attentes en termes de fonctionnalités. Bien sûr, ces « cibles » sont toujours lacunaires car le but d'une telle ouverture des données est aussi de toucher des publics éloignés de la culture, que nous connaissons donc mal et qui pour autant ne doivent pas être omis dans nos prospections et dans la projection de ce que nous comptons créer.

Pour que le MN puisse confirmer son statut nouvellement acquis de promoteur d'un lien renouvelé avec le public axé sur l'usage des technologies numériques de dernière génération pour rendre accessible et valoriser objets, savoirs et savoir-faire, le *crowdsourcing* est apparu comme une suite intéressante à considérer, les avantages comme les inconvénients d'une telle démarche étant à évaluer. Quel message peut véhiculer le MN en s'ouvrant à la collaboration d'amateurs, de non-experts pour documenter ses collections, participer au rayonnement de ses fonds et en faciliter l'accès ? Qui parmi toutes les Personas que nous avons relevées précédemment pourraient s'investir dans ce projet, versant conception comme versant contribution ? Et enfin, qui sont ces « autres » personnes vers qui nous devons aussi nous tourner pour nous assurer une communauté suffisamment étendue et diversifiée dans ses aptitudes et ses intérêts ?

Répondre à ces interrogations est essentiel car en fonction des personnes que nous

¹⁰⁹ Le détail de toutes les Personas est disponible en Annexe II, p. 68.

pourrons toucher, le choix des activités à mettre en place se précisera, l'idée étant de pouvoir paramétrer l'offre, les outils et l'ergonomie aux inclinaisons et aptitudes du plus grand nombre¹¹⁰. Pour dessiner nos Personas, nous avons interrogé des conservateurs du MN ayant déjà réalisé des indexations et des transcriptions pour leurs recherches universitaires, des personnes externes à l'institution mais entretenant avec elle un lien fort, et enfin repris les profils que les différents porteurs de projet que nous avons rencontrés avaient repérés¹¹¹. Bien sûr, nous ne pouvons pas présager de quelles catégories socioprofessionnelles et classes d'âge émergeront nos super-contributeurs - ferments d'un projet collaboratif et garants de sa progression qualitative et quantitative -, et ceux qui participeront moins mais seront tout aussi précieux pour le rayonnement du projet, mais ce travail de recueil est utile pour cibler les attentes et les erreurs à éviter. Nous verrons que les solutions collaboratives envisagées pour le MN laissent transparaître d'emblée les personnes les plus susceptibles de participer.

2. Créer un projet collaboratif : la réalité du terrain

2.1. Un projet au long cours et des modalités de contribution évolutives

Durant notre présence au MN, nous avons pu prendre la mesure de la concrétisation au quotidien des missions que s'est fixée l'institution, notamment ce qui a trait à l'ameublement des lieux de gouvernance et des problématiques que cela recouvre (utiliser des sièges ou des bureaux datant du règne de Louis XV quotidiennement suppose des partis pris en termes de conservation et de restauration et aussi une grande disponibilité des ateliers pour intervenir rapidement), à la logistique (organiser le transport des biens d'un ministère à un autre pèse lourd dans l'affectation des budgets et du personnel) et à l'informatisation des données afférentes aux collections susceptibles d'être utiles à tous les corps de métiers qui interagissent au MN. Si la mise en ligne des collections avait pour vocation principale de donner à voir la richesse du patrimoine au grand public, très vite le catalogue est aussi devenu un outil indispensable pour les inspecteurs, les documentalistes, la régie... D'où la volonté première de développer des fonctionnalités et des activités autour du MN/Lab afin de permettre aux usagers de trouver aisément ce qu'ils recherchent et surtout qu'ils bénéficient d'une information de qualité sur les objets grâce à des notices constamment enrichies.

Pour que la recherche dans le catalogue des collections soit plus pertinente et efficace, il faut que les objets soient mieux référencés. Mieux référencer, cela signifie enrichir la

¹¹⁰ Les profils imaginés des futurs contributeurs au projet du MN figurent en Annexe II, p. 74.

¹¹¹ Voir entretiens en Annexe II, p. 51-61.

description des biens par l'ajout de mots-clés par les usagers afin de caractériser de manière plus fine les œuvres grâce aux connaissances de chacun, depuis l'ébéniste qui indiquera la provenance probable de certaines essences de bois utilisées en marqueterie pour une époque donnée, jusqu'au botaniste-amateur qui identifiera les espèces végétales représentées dans les bordures d'une tenture et explicitera le message que souhaitait véhiculer l'artiste par leur recours. Ces ajouts, en fonction de leur acceptation par les modérateurs, de leur récurrence (qui peut traduire un intérêt prépondérant des internautes pour un domaine précis), de leur pertinence documentaire, peuvent permettre *in fine* la création de nouveaux critères de recherche, l'affinement des typologies existantes (types d'objets, auteurs, époque, style, manufacture et atelier) et l'usage de vocabulaires plus en adéquation avec les évolutions linguistiques contemporaines et la navigation sur le Web. Ces derniers seront certes moins spécialisés pour qualifier un meuble ou une tenture mais aideront néanmoins les néophytes à retrouver les objets qu'ils ont pu observer en contexte de visite par exemple, quitte ensuite à insister visuellement dans la notice-objet sur les termes adéquats utilisés par les professionnels, un lien vers le Portail des savoir-faire pouvant être plus qu'opportun¹¹². Pour le versant pratique de cette indexation libre, nous pouvons décider de rendre visible à tous les tags joints aux œuvres. Pour cela, nous créons un espace dédié à l'ajout de ces mots-clés sur chaque notice-objet, la possibilité d'en ajouter d'autres ou d'en retirer lorsqu'ils comportent des fautes d'orthographe ou n'induisent pas une meilleure description de l'objet, serait contingentée par la création d'un compte utilisateur (palier de sécurité contre les comportements malveillants). Si la modération des tags relève de la responsabilité des contributeurs, l'équipe-projet a néanmoins toute autorité pour retirer les tags jugés peu appropriés à la qualité informationnelle escomptée en recourant à ce type d'activité. Les tags agréés seront ensuite moissonnés par le moteur de recherche et deviendront des entités interrogeables directement en cliquant sur un tag. Une autre possibilité s'inscrirait dans le cadre plus contraint de *CollectionConnexion* où le détenteur d'un compte pourrait tagger les œuvres qu'il a sélectionnées pour figurer dans sa galerie personnelle, ces tags servant aussi bien à organiser sa collection, que de chemin d'accès pour les autres internautes lorsqu'elle est publique si les mots-clés sont moissonnés par l'interface. Les collections très bien décrites, illustrant une thématique ciblée, intelligemment articulées pourront être mises en exergue par les administrateurs du MN/Lab sur la page d'accueil du site afin de promouvoir le regard du public sur les collections.

¹¹² Au sujet du Portail des savoir-faire, voir la Partie III, sous-partie 2.2.

Également en lien avec le catalogue des collections, une deuxième hypothèse d'activité participative a été envisagée : elle a trait aux anciens numéros d'inventaires que certains objets portent. Par exemple, le tabouret en X de style Louis XIV époque Second Empire a pour numéro d'inventaire GMT 1487-002 utilisé couramment aujourd'hui pour y faire référence mais si nous examinons le revers de son armature, d'autres numéros sont inscrits et barrés, 20 378 TU et 3049¹¹³. Ces deux numéros témoignent de l'affectation de ce tabouret à plusieurs résidences (en l'occurrence aux Tuileries grâce à l'indicatif TU, et un autre lieu à déterminer pour le second numéro) car lorsqu'un bien servait à meubler un palais, à son entrée, un numéro lui était attribué et une courte description figurait dans le registre d'inventaire de l'année en cours (par exemple le registre d'inventaire des Tuileries pour l'année 1833, consultable aux Archives Nationales, côte AJ/9/171¹¹⁴). Si ce meuble venait à changer de lieu d'usage, un nouveau numéro pouvait lui être alloué et ainsi de suite. Parfois certains objets présentent plusieurs marques qui nous permettent de retracer leur histoire au cours du temps, qui les utilisèrent pour meubler leurs demeures et à quels moments ils ne furent plus au goût des gouvernants. Pouvoir rattacher ces anciens numéros reportés sur les notices des objets du catalogue MN/Lab à des palais et des châteaux pour une époque donnée grâce à la consultation de sources d'archives serait une réelle plus-value pour la documentation des collections du MN. Et faire exécuter cette tâche de recherche dans les fonds d'archives que nous avons préalablement relevés aux AN¹¹⁵ par des bénévoles semblait de prime abord correspondre à notre ambition de confronter davantage le public à l'héritage historique dont les collections du MN sont le témoignage, à cette œuvre de documentation s'adjoignant de plus une dimension éducative via cette familiarisation du public à la recherche en archives. Mais pratiquement, plusieurs écueils sont apparus au fil des tests que nous avons menés : si nous possédons les anciens numéros d'inventaire et des indicatifs de lieux, nous ne savons pas à quelle période le numéro a été affecté et il faut donc parcourir pour notre tabouret en X tous les registres d'inventaire des Tuileries où le numéro 20 378 apparaît. En effet, lorsque les registres d'inventaire de palais étaient rédigés, les objets présents dans une

¹¹³ Adresse du MN/ Lab: <https://collection.mobilienational.culture.gouv.fr/objet/GMT-1487-002>.

¹¹⁴ Archives nationales, lien vers le volume 3 du registre d'inventaire des Tuileries pour l'année 1833, localisée à la cote AJ/19/171 : www.siv.archives-nationales.culture.gouv.fr/siv-mediaFRAN_IR_027873_d_1_3_5_30_FRAN_0030_020816_L.

¹¹⁵ Un tableur Excel répertorie l'intégralité des registres d'inventaire de palais de la série AJ/19 (Compiègne, Tuileries, Saint-Cloud, Château de Versailles, Grand Trianon, Rambouillet, Élysée), et des séries O1 (Garde-Meuble de la Couronne sous l'Ancien Régime), O2 (Maison de l'Empereur) et O3 (Maison du Roi sous la Restauration). Certains fonds sont numérisés et consultables en ligne tandis que d'autres doivent être dépouillés en salle de lecture aux AN.

même pièce étaient enregistrés les uns à la suite des autres sans que le numéro affecté ne traduise leur présence ancienne ou leur arrivée récente dans les murs. C'est une recherche fastidieuse, surtout que tous les registres ne sont pas numérisés et accessibles en ligne, et parfois vaine car au numéro 20 378, nous n'avons par exemple jamais retrouvé notre tabouret. Il est donc impossible de prendre le problème dans ce sens, d'où la troisième solution retenue.

Cette troisième proposition, suggérée par Karl Pineau, doctorant en sciences de l'information et de la communication, nous est apparue plus tardivement, non en raison d'une mise en œuvre techniquement plus complexe ou d'un intérêt moindre pour le public, mais parce que les retombées en termes de recherche, de documentation et de développements ultérieurs ne pourront être véritablement palpables que lorsque l'entièreté du corpus envisagé sera traitée et que les chercheurs auront une base de données complète et normalisée pour permettre sa diffusion, son interopérabilité et sa pérennité. Nous reviendrons sur cette solution plus en détail car c'est celle qui a été la plus travaillée afin de favoriser sa mise en œuvre future par les personnes qui prendront notre suite. Cette activité orientée transcription n'aura pas forcément l'audience des deux premières options mais elle aura l'avantage d'ouvrir des pistes de recherche inédites du fait d'une mise à disposition de fonds d'archives encore peu exploités, promesse de découvertes et de collaborations extrêmement enrichissantes.

2.2. Des contraintes humaines et financières à la mise en œuvre du module participatif : ce qu'il advient aujourd'hui.

Privilégier la troisième solution participative plutôt que les deux autres est un parti pris assumé et mesuré à l'aune des réalités intrinsèques à l'institution, que ce soit en termes de fonctionnement, des relations interprofessionnelles, des chantiers en cours et d'une actualité pour le moins sensible et compromettante depuis la remise du rapport de la Cour des comptes en février 2019. En réaction aux accusations proférées par les magistrats de la rue de Cambon à propos d'une « *gestion du personnel sclérosée, de réserves pléthoriques et de créations trop peu valorisées* »¹¹⁶, le directeur du MN, Hervé Lemoine, a présenté aux agents de l'établissement un schéma directeur visant à redéfinir les missions premières du MN (à défaut de transformer profondément ou de recomposer l'ensemble comme le préconisait le rapport)

¹¹⁶ Voir le rapport public annuel 2019 de la Cour des comptes, tome 1, p. 245 à 274 « Le Mobilier nationale et les Manufactures nationales des Gobelins, de Beauvais et de la Savonnerie : une institution à bout de souffle », <https://www.ccomptes.fr/system/files/2019-02/07-mobilier-national-manufactures-nationales-Gobelins-Beauvais-la-Savonnerie-Tome-1.pdf>.

tout en conservant ce qui fait la spécificité et l'identité du MN et en poursuivant la dynamique d'ouverture entamée depuis deux ans. Pour être en conformité avec cette politique énoncée début mai et suivre le calendrier des développements prévus par l'équipe EIG, nous avons dû adapter nos propositions, et surtout entériner le fait qu'elles ne pourraient pas voir le jour en 2019, le collaboratif ne faisant pas partie des priorités du MN. En effet, la Direction a décidé de mettre à profit les derniers mois de présence de Ned Baldessin et Laurie Chapotte pour accroître les fonctionnalités du catalogue des collections, notamment en développant *CollectionConnexion* qui sera officiellement disponible en septembre prochain, et le Portail des savoir-faire qui est une création *ex nihilo*, c'est-à-dire que les contenus n'existant pas (à la différence du catalogue des collections MN/Lab dont le contenu repose sur une extraction de la base de données et de gestion des collections interne SCOM), il va falloir les créer intégralement, définir une ligne éditoriale pour leur donner une cohérence et une orientation, et écrire les articles en collectant les savoirs des artisans actifs dans les ateliers de restauration et de création du MN. Ce Portail des savoir-faire est un projet très important pour nombre d'acteurs de l'institution et issus du monde des métiers d'art car la crainte de l'érosion, voire de la perte de techniques ancestrales (et des formations dédiées à leur enseignement car trop souvent phagocytées par le système LMD et les universités), est une problématique actuelle qui mobilise tout cet écosystème qui œuvre à leur mise en valeur et leur transmission.

À ces contingences organisationnelles s'ajoutent de réelles réticences de la part du personnel à l'égard du collaboratif, ces derniers, déjà très accaparés par leurs missions, n'imaginant pas devoir consacrer du temps à une interaction plus poussée avec le public, que ce soit en termes de modération ou de validation des contributions. Tout ceci nous a conduits à réévaluer les possibles activités participatives à mettre en place, déjà écrémées pour des raisons humaines et budgétaires. Les entretiens menés nous ont aidés à faire ce tri et à percevoir à quel point certains outils et certaines procédures de vérification représentent un travail trop conséquent par rapport aux disponibilités humaines et financières du MN.

Pour revenir brièvement sur les raisons nous ayant conduit à écarter le travail sur les anciens numéros d'inventaires, rappelons que l'expérience de la consultation et du dépouillement d'archives qu'il requiert, tout comme le caractère fastidieux de la recherche et l'exigence de rigueur et d'esprit d'analyse dans la lecture des sources, font que la communauté de contributeurs aurait forcément été restreinte, certainement soudée et fidèle car l'investissement consenti demande un intérêt très fort pour la problématique et une capacité à

dialoguer avec ses pairs, mais finalement peu ouverte au grand public (ce qui ne correspond pas aux attentes de la Direction). Scientifiquement parlant, cette proposition est très intéressante car reconstituer la logique qui présida à l'attribution des numéros d'inventaires et à l'inscription des biens dans les registres du Garde-Meuble et des palais est un exercice « *pas du tout valorisant mais fondamental pour l'histoire de l'art et des collections* » selon les mots de Pascal Bertrand, la vocation même d'un garde-meuble où c'est le principe de nécessité qui l'emporte sur le principe de conservation des biens culturels, ceux-ci étant destinés à être utilisés, déplacés, parfois malmenés, réunis en ensemble dans un lieu puis dispersés ailleurs, et non exposés en une place et muséalisés, fait tout l'intérêt d'une telle étude. L'atout de cette proposition était que cette communauté d'intérêt, formée de personnes très motivées et forcément averties des enjeux de ce travail, aurait été autonome, en capacité d'interagir, de se corriger et de s'autogérer elle-même sans nécessiter une intervention très poussée de l'équipe-projet et la nomination d'une personne chargée de la soutenir. Mais l'inconvénient de la tâche, telle que l'avons dans un premier temps conçue, est sa complexité qui risque de ne pas brasser un public autre que celui qui est déjà en lien étroit avec le MN.

Le *tagging* semble a priori moins problématique mais implique toutefois une intervention plus soutenue de l'équipe-projet pour assurer le contrôle régulier de la qualité des mots-clés ajoutés, censés permettre un meilleur référencement des biens et donc leur découverte plus aisée. Si nous souhaitions initier ce type d'activité, il faudrait mieux connaître ceux qui viennent consulter le site MN/Lab et pourraient être intéressés par ce type d'interaction : les chiffres de fréquentation collectés à l'issue de la mise en ligne des collections donnent déjà des indications et l'ouverture prochaine de *CollectionConnexion* pourra nous aider à statuer sur la pertinence à généraliser le recours au *tagging* pour améliorer un pan de la documentation des objets. En effet, de l'observation des typologies d'objets qui sont préférentiellement sauvegardées, thématiques et décrites (ajout d'un commentaire introductif et de mots-clés) par les internautes sur *CollectionConnexion* et du constat que des époques ou des artistes en particulier sont délaissés, naissent les orientations à donner pour une meilleure valorisation de ces « oubliés ». Des sélections effectuées par les documentalistes et les inspecteurs peuvent remédier à ces oubliés et inciter les usagers à leur prêter attention, se renseigner puis partager leurs informations via des mots-clés pour que d'autres les découvrent par sérendipité, pourrait en être une. Le travail de post-traitement est aussi conséquent pour l'équipe-projet qui se devra de suivre les itérations des taggeurs et d'impulser la dynamique adéquate pour améliorer le catalogue des collections, la modération des annotations pouvant

être déléguée aux contributeurs ayant créé un compte et accepté la charte de participation.

Pour le projet de transcription des registres d'entrées et de sorties du Garde-Meuble, de fabrication des manufactures de Beauvais, des Gobelins et de la Savonnerie, des meubles confectionnés, des prêts de tapisseries, des plâtres et dessins, des ateliers de restauration dans le but de construire une base de données où nous disposerions de toutes les informations relatives aux objets (création, usages, restaurations), ce sont les choix opérés pour le développement tant de l'interface de travail, les outils d'aide à la contribution (tutoriels, FAQ), les modalités d'échanges, que les ressorts cognitifs sollicités (récompenses, ludisme) qui obligeront l'équipe-projet à intervenir plus ou moins régulièrement et à quel degré. Toutefois, pour que la base de données puisse être perçue comme une source fiable et réutilisable pour la recherche, il faut que les transcriptions soient soumises à évaluation. Cette évaluation, qu'elle soit automatique, semi-automatique ou manuelle, réalisée pour partie par les contributeurs ou entièrement prise en charge par l'équipe, doit être robuste avant toute chose. L'avantage des solutions logicielles clé en main comme *ScribeAPI* est que ces modalités sont déjà prédéfinies dans le package, mais encore faut-il que cela convienne aux spécificités du projet et aux attentes des administrateurs. C'est pourquoi rattacher notre ambition à celle que poursuit Guillaume Raschia au sujet de la création d'une plate-forme numérique de production participative à destination des projets de recherche en Sciences Humaines et Sociales, Arts et Lettres, intitulée 3Puma, nous offrirait un terrain privilégié pour discuter et expérimenter des procédures de validation les plus autonomes possibles afin d'alléger la charge de travail des institutions souhaitant recourir au *crowdsourcing* tout en n'ayant pas forcément le personnel disponible pour en assurer le suivi au long cours. Nous allons à présent revenir sur les circonstances nous ayant conduits vers le projet 3Puma, ses promesses comme la réalité d'un tel partenariat et ce que cela pourrait signifier pour le développement du participatif au MN.

3. De la prospection à la concrétisation du module de transcription collaborative : des débuts prometteurs

3.1. Une communauté professionnelle porteuse des valeurs du crowdsourcing

La rencontre de porteurs de projets dans le cadre du stage a été une réelle plus-value pour notre découverte du phénomène et la prise de conscience des enjeux et des impératifs que requiert la mise en œuvre d'un projet de *crowdsourcing* dans le cadre des GLAMs. Il est

d'ailleurs intéressant de remarquer que la logique du dialogue et du partage qui prévaut dans tout projet de *crowdsourcing* réussi s'applique aussi bien aux contributeurs qu'aux concepteurs qui travaillent aussi ensemble pour que le procédé devienne une méthode de recherche à part entière, aussi fiable qu'une autre, avec tout son arsenal d'outils propres. Les personnes contactées ont aussi bien transmis leur expérience du participatif que donné des conseils pratiques et des orientations pour notre projet de reconstitution des historiques des objets dont l'intérêt a été reconnu unanimement. Ainsi, Yves Carlier, conservateur général du département des ressources documentaires du château de Versailles, et ses collègues voient dans ce travail sur les numéros et les marques attribués au cours du temps aux objets conservés au MN le prolongement logique de leurs propres recherches à ce sujet et donc l'opportunité de développer plus encore le répertoire des marques qu'ils ont établi, de compléter l'historique de leurs propres objets et de relier plus étroitement encore objets et archives avec la mise à disposition de celles que détient le MN¹¹⁷.

Alors que nous avons des difficultés à conceptualiser l'activité qui nous permettrait d'exploiter au mieux les ressources dont nous sommes dépositaires tout en attirant suffisamment de bénévoles pour relever le défi imaginé, deux rencontres ont été éclairantes sur la conduite à observer pour répondre aux attentes des uns et aux contraintes qui sont les nôtres. Tout d'abord, Karl Pineau nous a conseillé de créer une base de données relationnelle qui aurait l'avantage de regrouper toutes les ressources que nous possédons à propos des objets du MN, à savoir les registres énumérés en 2.2. qui jusqu'ici n'ont été exploités que partiellement et jamais reliés les uns aux autres faute d'un outil facilitant le rapprochement. Or, pour répondre à des questions de recherche centrées sur l'histoire économique et sociale de l'art, nous avons besoin de rassembler tous ces documents dans un même espace pour pouvoir les traiter simultanément et reconstituer brique par brique l'imbrication des individus, des événements et des besoins présidant à la production de biens mobiliers. Ce qui implique dans un premier temps de découpler l'activité de transcription des registres des recherches futures afin que nous soyons sûrs que les vides apparents dans la vie des objets ne sont pas de notre ressort mais doivent être comblés en exploitant d'autres archives que les nôtres (l'adjonction de ressources extérieures à la base de données étant inscrit dans le processus participatif). En effet, l'intérêt de construire une base de données relationnelle est qu'elle peut sans cesse évoluer dans les liens entre entités et objets qu'elle permet et que si nous ne

¹¹⁷ Voir en Annexe II, p. 59, l'entretien avec Yves Carlier explicitant le fonctionnement de la base de données des collections du château et du domaine de Versailles, notamment la manière dont sont renseignés les anciens numéros d'inventaires et les marques des biens, et la construction d'une mise en regard entre archives et objets.

parvenons pas à relier des biens recensés dans les inventaires de palais avec nos propres collections, rien ne dit que les enrichissements documentaires à venir, grâce à la structuration que permet cet outil de traitement de l'information, ne pourront pas autoriser des éclaircissements au fur et à mesure de la progression même de notre connaissance et de l'exploitation de ces archives. La mise en forme des ressources, et éventuellement leur encodage, est donc une nécessité pour faciliter leur accès et leur interprétation.

Intervient alors une deuxième personne, Florence Clavaud, qui en plus d'accréditer le fait de commencer par l'action la plus simple, l'indexation des archives, pour amorcer un projet en histoire de l'art d'envergure, nous a indiqué comment construire un modèle de données robuste qui permettra de faire le va-et-vient entre les objets conservés actuellement au MN et les archives. Pour cela, il faut modéliser les objets, c'est-à-dire les définir et donc comprendre la structure qu'ils traduisent. Dans nos inventaires, nous avons les traces de la création des biens culturels (commande, concours), des modifications qu'ils ont subies au cours de leurs usages successifs (par exemple, retissage des textiles présentant les symboles de l'Empire pour qu'ils puissent être réutilisés par le Roi à la Restauration), de leurs envois de résidence en résidence pour le service de telle ou telle personne. Ce qui nous intéresse ici est de tracer tout ceci, de montrer l'imbrication des événements, des lieux, des personnes, des collectivités dans ces circulations et d'imaginer la manière dont les chercheurs ou les amateurs vont interroger ces registres, utiliser les informations qui vont être transcrites et modélisées via cette base de données. De ces agencements découleront les conditions de remplissage de la base et les interfaces pour ce faire. Si c'est le développeur qui sera chargé de gérer ces données en termes d'objets, de relations, de propriété, nous avons néanmoins besoin de définir en amont les systèmes de descriptions et les nomenclatures les plus adaptés à notre sujet. Pour ce faire, Florence Clavaud a recommandé l'usage du modèle conceptuel HADOC¹¹⁸ du ministère de la Culture et de la communication qui offre un cadre normatif adapté à la production de données culturelles et à leur diffusion via une harmonisation et un alignement des référentiels et des standards sur ceux utilisés à l'échelle mondiale. En assurant l'interopérabilité de nos données grâce à ce modèle, nous leur conférons une visibilité sur le web inédite qui permettra leur exploitation et leur incrémentation futures.

¹¹⁸ Le programme HADOC (Modèle harmonisé pour la production des données culturelles) a été lancé en 2008 par le MCC pour unifier la saisie, les formats et la granularité des données utilisées pour décrire une œuvre d'art et répondre aux besoins d'intégrations de ces dernières sur le web sémantique.

Enfin, une dernière rencontre est venue offrir des perspectives fort prometteuses pour la suite de notre projet : Guillaume Raschia, chercheur en sciences du numérique, n'a pas seulement été de bon conseil quant au choix des logiciels et des solutions techniques pour développer une activité participative autour de la transcription d'archives, mais il a proposé d'inclure le MN dans le consortium qu'il a fédéré autour de son projet de création d'une plateforme francophone destinée à accueillir des projets de *crowdsourcing* en lettres, arts et sciences humaines, possédant sa propre communauté de contributeurs fidèles et mobilisables immédiatement sur les projets, à l'image de ce que propose Zooniverse dans le monde anglo-saxon. Nous développerons ce point dans la partie suivante mais nous souhaitons auparavant revenir sur le projet *RECITAL* car l'analyse des choix opérés du point de vue des logiciels est intéressante pour nous afin de comprendre que si nous ambitionnons de documenter des collections par le recours au *crowdsourcing*, nous avons pour obligation de construire un système où l'informatique viendra faciliter l'action de l'humain et non s'y substituer, offrir des garanties en termes de qualité des données sans pour autant oublier que c'est l'humain qui donnera du sens à cette accumulation d'informations et pas la machine.

Ainsi, pour le projet *RECITAL*, Guillaume Raschia et son équipe ont choisi le framework ScribeAPI développé par Zooniverse et la NYPL Labs qui certes offre tout un outillage (briques logicielles en Open Access) pour la mise en œuvre d'une activité de transcription de manuscrits segmentée en micro-tâches (marquage des données à transcrire dans le document, transcription et vérification par les pairs) mais pose d'importants problèmes de maintenance et d'évolution des fonctionnalités. En effet, c'est une solution qui est aisément déployable pour un projet de dimension modeste, des fichiers de configuration en Json (texte structuré) permettent de définir les modalités de traitement des sources formant le cœur du projet sans avoir besoin de savoir coder, mais si nous souhaitons internationaliser l'initiative, c'est-à-dire traduire en français l'intégralité des chaînes de caractères, il faut modifier le code. Or celui-ci est fait de telle manière qu'il est compliqué d'ajouter des micro-fonctionnalités localement sans perturber d'autres éléments à l'autre bout de la plateforme. L'amélioration des possibilités du logiciel pour convenir aux objectifs fixés par *RECITAL* prend alors beaucoup de temps, demande des compétences élevées et n'a de surcroît pas entièrement satisfait l'équipe qui n'a pu aller au bout de son souhait d'étudier de manière approfondie les phénomènes de redondance, d'allocations de documents et de tâches en fonction du profil des contributeurs afin d'améliorer la qualité des données produites. De plus, l'équipe n'a pu procéder à la validation des transcriptions qu'une fois l'entièreté du corpus transcrit, donc a

posteriori et en dehors de la plateforme, ce qui n'était pas prévu initialement et n'est pas idéal pour évaluer la qualité de la production des contributeurs, mais l'usage de ScribeAPI ne permettait pas de procéder autrement, modifier cet état de fait étant trop lourd pour une équipe disposant de moyens humains et financiers modestes. D'où cette volonté du Laboratoire des Sciences du numérique de Nantes de fédérer un groupe de réflexion autour de ces problématiques et d'ouvrir de nouvelles perspectives techniques qui répondront mieux aux attentes de la recherche envers le crowdsourcing pour produire de nouvelles connaissances.

3.2. Nouer des partenariats et lancer des expérimentations : la manufacture de Beauvais et le projet Arachné

Le MN a été invité à rejoindre le projet 3Puma coordonné par Guillaume Raschia¹¹⁹ en tant que testeur de la plateforme sur ses propres corpus documentaires et aussi pour que le MN propose des fonctionnalités à développer afin d'améliorer le processus de participation. Cette opportunité nous a semblé plus qu'intéressante à saisir car en nous insérant dans des consortiums de recherche et en répondant à des appels à projets comme celui de l'ANR « *Sciences ouvertes* » sur « *le développement de plateformes de services* », c'est bénéficier d'une offre de service pérenne, intégrée dans l'écosystème des institutions culturelles nationales, disposant de moyens humains et financiers bien plus conséquents que ce que nous pourrions obtenir nous-mêmes pour créer notre plateforme, avec des solutions de maintien prévues en amont grâce à la mutualisation des coûts. Tous les objectifs que poursuit la plateforme 3Puma entrent en concordance avec les retours d'expériences que nous avons recueillis, à savoir qu'initier individuellement un projet de *crowdsourcing* est très complexe car nous n'avons jamais la certitude qu'il rencontrera son public et que nous pourrions mobiliser une équipe suffisamment longtemps pour accompagner la communauté sur le long terme et assurer la pérennité des données pour des exploitations futures. Et justement à propos des questions de maintien de la solution, conscients de la méfiance que peuvent susciter les projets qui éclosent dans le cadre des appels à projets émis par l'ANR, notamment parce que celle-ci n'octroie des financements que pour une courte période (un an et demi à deux ans), Guillaume Raschia et son équipe ont trouvé en la MSH Ange Guépin, opérateur du dispositif Box Huma-Num¹²⁰, un allié de choix qui s'est engagé à assurer les coûts de maintenance

¹¹⁹ Voir en Annexe III, p. 92, le détail du projet proposé en réponse à l'appel lancé par l'ANR en mai 2019.

¹²⁰ La Huma-Num Box est un dispositif de stockage distribué en réseau permettant d'accueillir des jeux de données de grande taille, souvent dites « tièdes » ou « froides » car peu d'accès en écriture ou en lecture seront effectués durant la vie de ces données, tout en assurant une conservation fiable durant cinq à dix ans de ces

technique, d'infrastructure et d'animation bien après la fin du financement alloué par l'ANR. Quant à la question d'ouverture et de réutilisation des données produites par les projets hébergés sur la plateforme 3Puma, l'adoption des principes FAIR¹²¹, la standardisation au format XML/Dublin Core et le moissonnage OAI-PMH ont pour vocation de rendre possibles ces deux aspects sensibles et fondamentaux pour ce projet de sciences des données et assurer la promotion du *crowdsourcing* comme une méthode reconnue de recherche scientifique.

Si les fonctionnalités que l'équipe-projet 3Puma escompte développer en priorité rencontrent nos propres attentes, - possibilité d'annoter et de transcrire des documents, de créer des comptes personnels pour les bénévoles, de disposer d'indicateurs d'activité et de progression du projet, de protocoles visant à assurer l'interopérabilité et la réutilisation des données pour des exploitations ultérieures -, la taille de notre corpus (très conséquent si nous prenons en compte l'intégralité des registres numérisés mentionnés plus haut et ceux des AN) et la mobilisation des équipes concernées par notre projet sur d'autres sujets prioritaires dans la politique d'ouverture du MN sur le Web, font que nous n'étions pas assez avancés pour répondre à l'offre de G. Raschia dès mai 2019. Néanmoins, compte tenu de l'intérêt et du caractère étayé de la proposition, nous avons suggéré à Pascal Bertrand, porteur du projet *Arachné* sur la manufacture de tapisseries de Beauvais, de rejoindre le consortium 3Puma pour que le travail qu'il mène depuis quatre ans avec le Getty Research Institute et ses élèves de Master puisse trouver une valorisation à la mesure de son intérêt pour l'Histoire de l'art.

Souhaitant donner un nouvel essor à la recherche portant sur la manufacture de tapisseries de Beauvais mais en atténuant la dimension artistique, largement traitée auparavant, pour se concentrer aussi sur une approche socio-économique et industrielle de la question, Pascal Bertrand et Charissa Bremer-David ont commencé à dépouiller les registres de fabrication de la manufacture sur une période de 70 ans. Ces livres de compte fournissent des informations sur le métier de lissiers, les réseaux professionnels qui s'organisent autour de la confection d'une tenture, sur la clientèle et le goût qui prévaut à une époque donnée. Le recoupement de ces données avec celles qu'ont compilées d'autres chercheurs depuis le XIXe siècle permet d'aborder « *l'histoire externe des œuvres, celle qui commence après leur livraison et qui se*

matériaux de recherche à consonance patrimoniale. Ce service se situe entre les services d'Huma-Num à destination des utilisateurs finaux (Sharedocs et NAKALA) et le service d'archivage à long terme.

¹²¹ Le FAIR data est un écosystème dont l'objectif est de s'assurer que les données produites sont trouvables, accessibles, interopérables et réutilisables grâce à une construction, un stockage et une présentation définis en amont afin d'être en mesure de répondre à une croissance exponentielle des données et à la nécessité de disposer d'outils pour mieux les structurer et les réutiliser.

rattache à l'histoire des collections »¹²². Suivre le parcours de ces œuvres n'est pas aisé car elles ont subi nombre d'avatars et de modifications (dans leurs dimensions et leurs motifs) qui compliquent singulièrement le travail de correspondance entre les tapisseries décrites dans les registres et celles qui sont désormais conservées par les musées (surtout américains) ou sont en mains privées. Il faut donc mener des recherches dans les nombreux fonds documentaires et iconographiques anciens pour pister ces tapisseries de demeure en galeries d'exposition et suivre leur réception au fil du temps, la considération qui leur était accordée et la nature du dialogue qui était instauré entre ces œuvres et le mobilier des intérieurs dans lesquels elles étaient installées. Cette partie du projet *Arachné* est très conséquente, d'où l'idée de faire appel à un groupe plus large pour dépouiller ces ressources et ces registres, créer des liens et le besoin identifié de disposer d'une plateforme pour mutualiser et organiser la consultation des fonds, mettre à disposition de tous les découvertes pour faire progresser la réflexion et donner enfin une envergure et une diffusion à la mesure de l'ambition de cette question de recherche passionnante. Ce projet, en plus de posséder tous les atouts pour fédérer une communauté d'intérêt forte d'amateurs et d'experts, va nourrir les questions de recherche en sciences des données et de l'information auxquelles tente de répondre le Laboratoire de Nantes à travers le développement de la plateforme 3Puma. Enfin, le projet de Pascal Bertrand et Charissa Bremer-David, parce qu'il entretient des connexions évidentes avec ce que nous escomptions mettre en œuvre, servira de préfiguration pour nos propres objectifs et nous aidera à faire les bons choix au moment dit. Nous attendons donc de connaître les projets retenus par l'ANR et les premiers retours expérimentaux si c'est la plateforme 3Puma qui a les moyens d'émerger et de donner au *crowdsourcing* un nouvel essor en France.

3.3. Les modalités de mise en œuvre du projet de transcription : présent et avenir

En attendant cette échéance, et en cas de non recevabilité de 3Puma, nous avons poursuivi le travail de préfiguration et de préparation du projet participatif du MN afin de faciliter au mieux le travail de l'équipe qui assurera sa mise en œuvre technique et opérationnelle. Pour que la base de données que nous souhaitons créer soit réellement un outil permettant d'accroître la référentialité de nos archives, c'est-à-dire leur permanence et leur accessibilité, nous avons commencé par décrire précisément nos registres, leurs caractéristiques avant de les découper en unités documentaires qui permettront de trier et filtrer les informations selon

¹²² Annexe III, p. 104 : Description du projet *Arachné* par Pascal Bertrand et Charissa Bremer-David.

les besoins du visiteur¹²³. Une première unité documentaire est le registre dont le titre, les bornes chronologiques, le lieu de conservation, le nombre de pages et le permalien sont à spécifier. Le contributeur choisit un type de registre sur lequel travailler et quand il accède au formulaire dudit registre, les champs afférents à cette première unité documentaire seront pré-remplis. Ensuite vient la deuxième unité documentaire qui est la page de ce registre. Dans cette partie du formulaire figurera le numéro de la page en cours de traitement et une description matérielle de la page que le contributeur doit effectuer : mention de son état de conservation, de la présence de déchirures, de taches, de manques, d'une encre passée, une graphie complexe à déchiffrer, d'un papier de moindre qualité, bref, tout ce qui pourrait expliquer des difficultés de lecture du contenu et d'interprétation. La saisie de ces informations peut se faire sous la forme d'un menu déroulant où figurent tous ces écueils, le bénévole cochant les cases utiles, ceci afin de permettre une harmonisation maximale du vocabulaire utilisé (et pour rendre plus aisé le traitement ultérieur des données). Enfin, la troisième unité documentaire identifiée est l'objet. La transcription se fera objet par objet, pour chaque objet, des champs dédiés étant ouverts selon ses caractéristiques. En effet, selon le type de registres, ce ne sont pas les mêmes informations qui sont renseignées et deux choix se présentent à nous : soit nous reprenons l'organisation première de chaque registre reproduite en annexe IV, si le contributeur choisit de traiter un registre de palais, le formulaire avec les champs renseignés originellement dans ce registre apparaissent quand le contributeur choisit de traiter ce type de registre, et ainsi de suite s'il opte pour un registre de confection de tapis, un inventaire des dessins, plâtres et ciseaux, un registre de prêts de tapisseries, etc. Une seconde option consiste à créer un formulaire unique de saisie pour tous les types de registres recensés au MN, certaines informations étant redondantes d'un registre à un autre mais d'autres seront forcément laissées de côté car très spécifiques à un type d'objets qui ne sont pas généralisables à tous¹²⁴. Les champs qui apparaîtraient en ce cas seraient relatifs à la désignation de l'objet, son auteur (artiste, artisan), sa description, ses matériaux constitutifs, ses dimensions, sa destination ou sa localisation au moment du récolement et sa valeur monétaire. L'inconvénient de cette solution est de construire une base de données qui ne répertoriera pas l'entièreté des données dès maintenant et ne donnera donc pas une vue

¹²³ CHUPIN Lisa, *Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés*, 2017, 220 p., Thèse de doctorat en Sciences de l'information et de la communication, Paris, CNAM, sous la direction de Manuel Zacklad., p. 130.

¹²⁴ Voir en Annexe II, p. 79 les modalités de remplissage de ces champs tout en sachant qu'elles évolueront en fonction des orientations choisies, notamment pour ce qui a trait aux unités de mesure, aux convertisseurs à insérer et aux aides à mettre en place pour que le contributeur puisse intégrer les règles établies.

exhaustive de la richesse informationnelle détenue par ces registres aux chercheurs qui souhaiteront les exploiter et qui devront forcément revenir à la source primaire en cas de manque. L'idée de créer un espace de mise en valeur de nos corpus qui fasse référence pour aborder une histoire de l'art plus centrée sur les métiers, les artisans et les réseaux entourant la fabrication des objets d'ameublement, est alors non avenue car parcellaire et orientée dès sa constitution vers la simplification, simplification qui ne sert pas les objectifs scientifiques du projet. Pourquoi alors envisager cette solution ? Parce que son atout est qu'en dégageant des communs à tous ces registres, nous pouvons plus aisément structurer les données et les recouper, choisir un encodage et des standards qui enregistreront un maximum d'informations, éviteront de possibles difficultés de publication et faciliteront l'extraction des informations¹²⁵. Nous n'avons pas les moyens de prévoir l'exhaustivité des usages qui seront faits de notre base, aussi devons-nous tout mettre en œuvre pour que le travail des contributeurs soit d'excellente qualité, le choix d'une solution plutôt que de l'autre ouvrant sur des exploitations et des modes de contributions légèrement différentes. Conscients du caractère rébarbatif de ce travail de transcription, nous avons aussi souhaité laisser au contributeur un espace où il pourra partager le résultat de ses recherches personnelles à propos du sujet dont il est question dans l'archive en cours de transcription, par la citation d'ouvrages qu'il a consultés et surtout en rattachant à l'objet un visuel qui peut être celui disponible du MN/Lab quand l'objet est conservé au MN ou sur un autre catalogue de musée quand l'objet est conservé à l'étranger par exemple. Tout en amenant une dimension ludique, ce travail de correspondance entre l'archive et les œuvres actuellement conservées dans des institutions patrimoniales ou en mains privées est un pan fondamental de la recherche en histoire de l'art, tous les historiens passent par ce jeu de repérage et tout lien établi sera précieux pour l'utilisation future de notre base. Cette attribution comporte parfois un caractère incertain, notamment pour les textiles dont les dimensions peuvent avoir été modifiées au cours du temps, et les contributeurs auront alors la possibilité d'évaluer le degré de certitude de leur proposition, à l'image de ce que propose *Digital Mellini* : « *identifié, probablement identifié, œuvres liées, non identifié* » peuvent être les mentions retenues. Enfin, un espace de commentaires libres laissera la possibilité aux volontaires d'échanger autour de l'objet, de la transcription du registre, des difficultés rencontrées, des interrogations soulevées, etc.

¹²⁵ LABIA Camille, *Les humanités numériques au service d'un projet de recherche collaborative en archéologie*, septembre 2018, 70 p., Mémoire de stage, Paris, École du Louvre, sous la direction de Françoise Dalex et Christophe Leclercq, p. 21.

Pour être en adéquation avec la politique du MN, l'interface de transcription doit être conçue comme un espace de médiation et d'émulation et être pour les contributeurs une occasion privilégiée d'interagir avec les collections et de construire ensemble, avec le soutien de l'équipe-projet, des connaissances à leur endroit. Pour favoriser l'engagement et la fidélité des volontaires, et contrebalancer le cadre contraint de participation dû à la nature même de l'activité proposée, ces espaces de libre expression sont à notre sens essentiels et nous devons absolument soutenir l'expertise des contributeurs en leur confiant un statut particulier au sein du projet, que ce soit en tant que modérateurs ou vérificateurs, et en les encourageant à suggérer des ressources complémentaires à consulter et intégrer dans la base ou à co-créeer les projets de recherche qui pourraient naître de cette compilation inédite de ressources à propos des collections du MN. Nous avons ainsi songé à créer sur notre plateforme un espace « bac à sable » où les contributeurs pourraient se projeter vers une véritable analyse de la démarche à laquelle ils participent et des données qui en émanent, des historiens de l'art et chercheurs de tout bord pouvant se joindre à la réflexion dans une logique de collaboration à part entière.

Au moment où nous clôturons ce mémoire, nous ne pouvons nous prononcer plus avant sur la forme que prendra ce projet participatif et s'il deviendra une réalité. Mais nous avons posé les jalons essentiels et envisagé les options possibles à mettre en place en fonction des moyens qui seront donnés aux personnes qui en assureront la réalisation. Nous concluons cette partie en insistant sur l'opportunité qu'offre la création d'un outil participatif dans le contexte des GLAMs pour réfléchir au statut de l'histoire numérique qui, comme le souligne Pauline Moirez, « *n'est pas seulement une méthode nouvelle qui surfe sur la vague des nouvelles technologies numériques, un nouveau champ de l'histoire, mais participe de l'histoire en tant que telle et oblige à réfléchir sur les canaux de transmission des discours historiques au-delà du livre* »¹²⁶. Nous avons à disposition un autre levier pour toucher le public et l'opportunité de reposer la question de la place de l'historien dans l'écosystème du numérique qui n'a en réalité pas à se positionner en garant absolu de la « juste » diffusion de la connaissance scientifique face à une masse d'internautes en apparence indifférenciée mais au contraire doit accompagner au mieux le vœu de chacun de prendre part à une histoire numérique qui interroge les méthodes traditionnelles pour construire un nouvel univers de sens.

¹²⁶ MOIREZ Pauline, MOREUX Jean Philippe, JOSSE Isabelle, *État de l'art en matière de crowdsourcing dans les bibliothèques numériques*, Livrable L-4.3.1 du projet de R&D du FUI 12 pour la conception d'une plateforme collaborative de correction et d'enrichissement des documents numérisés, février 2013, 77 p., [en ligne], consulté le 03 février 2019, URL : http://www.bnf.fr/documents/crowdsourcing_rapport.pdf.

CONCLUSION

Les pratiques participatives, considérées comme un moyen de mettre à profit la connaissance et les compétences du plus grand nombre sur des sujets ciblés, s'inscrivent dans une histoire de la recherche scientifique dont la pertinence n'est plus démentie et constitue un champ expérimental pour penser et organiser le monde à l'heure du Web 2.0 plus qu'intéressant à investiguer et développer. La mise en œuvre de tels projets demande néanmoins une certaine prudence car de cette ouverture au plus grand nombre pour traiter et interpréter des corpus de données naît une incertitude quant à la qualité de l'information ainsi générée et le risque de détournements à des fins de nuisance pour les personnes, physiques comme morales, impliquées. Pour autant, l'imbrication de logiques sociales, politiques, économiques et éthiques qu'induit cette démarche est source d'une grande richesse car en mobilisant et en érigeant la diversité, la mémoire et le respect des opinions et des personnes comme ferments de la participation, elle constitue un espace plus que fécond pour solutionner des modes de fonctionnements qui n'ont plus leur place aujourd'hui dans notre monde et aussi apaiser le rapport au passé des jeunes générations qui ont là une occasion de se réapproprier une partie de leur identité à travers le témoignage de leurs aînés.

Associer les valeurs que porte le participatif à celles que le Mobilier national s'efforce de véhiculer à travers les différentes modalités d'ouverture de ses murs et de ses collections au plus grand nombre est tout à fait cohérent, tout comme l'incitation faite au public d'interagir directement avec l'institution pour enfin découvrir ce qu'impliquent les missions qu'elle poursuit au quotidien en termes de savoirs et de savoir-faire de la part des agents qui y travaillent. Pour le Mobilier national, mettre en ligne ses collections, créer un Portail des savoir-faire et peut-être bientôt initier tout un travail de recherche autour des archives et de l'histoire des objets, est certes un moyen de s'insérer dans l'écosystème des humanités numériques dont il était éloigné, mais surtout de bousculer les représentations qui lui étaient attachées, à savoir un lieu destiné au service des élites et très renfermé sur lui-même.

Les points de discussion relatifs à la concrétisation d'un projet participatif sont, nous l'avons vu, nombreux, que ce soit en termes de conduite de projet au long cours, de réflexion

sur les objectifs qu'il faut impérativement atteindre et sur la manière dont les programmes vont se maintenir au fil du temps, se renouveler, s'améliorer, avoir la capacité de fédérer d'autres publics en s'adaptant à la fois aux centres d'intérêts des internautes et aux aspects que l'institution souhaite mettre en exergue par cet appel à la contribution. Il n'en demeure pas moins que la question de la création d'une communauté de collaborateurs reste centrale dans tout projet collaboratif et que cela demande du temps pour que celle-ci émerge, se fortifie et fonctionne en autonomie. De même, si mettre au point le scénario irriguant le projet et les outils les plus idoines pour le concrétiser demandent un travail bien en amont, il demeure crucial de ne pas omettre l'investissement conséquent et quotidien que requièrent les activités de modération, de vérification et de validation des contenus. C'est pourquoi l'idée de développer une plateforme comme 3Puma est tout à fait intéressante car elle permet aussi bien de répondre à nombre des difficultés susmentionnées et aussi de créer un environnement de travail satisfaisant autant les attentes des porteurs de projet (production de données qualitatives, contributeurs mobilisables instantanément, validation semi-automatisée) que celles des bénévoles (variété des sujets, convivialité, partage, reconnaissance).

Pour ouvrir plus largement le sujet, l'avènement de la collaboration sur le web et la création d'environnements propices à la redocumentarisation du patrimoine tendent à transformer le dispositif de médiation documentaire et à mieux surmonter les différences de représentations entre usagers et professionnels de l'information et des musées. Les outils de recherche, en s'appuyant sur des concepts qui émanent de la « foule » gagnent en pertinence en évoluant parallèlement au langage et aux usages des internautes. Nous pouvons percevoir ce qui questionne, alerte, mobilise, ce qui permet de dessiner les contours d'un paysage culturel et mémoriel renouvelé en phase avec les progrès technologiques. Il n'est pas anodin d'imaginer qu'un jour, l'intelligence artificielle parviendra à réaliser ce travail d'indexation, de transcription, de recirculation de la donnée sans passer par la main humaine. Mais pour atteindre un tel développement du Machine Learning, il faudra encore des années et des milliards d'éditions manuscrites pour que la machine apprenne suffisamment pour s'auto-corriger d'elle-même, rechercher des informations en plein texte dans les documents manuscrits, suggérer des mots-clés pour annoter et caractériser les images... L'intervention humaine demeure encore incontournable pour fournir une capacité de compréhension et une mémoire à nos ordinateurs.

BIBLIOGRAPHIE

Ouvrages-monographies, articles de périodiques, travaux universitaires, communications à congrès

ALAM Sultana Lubna, CAMPBELL John, « Crowdsourcing motivations in a not-for-profit GLAM context: The Australian Newspapers Digitisation Program », in *23rd Australasian Conference of Information Systems*. Geelong, 3 décembre 2012, p. 1-11.

ALAN Irwin, *Citizen Science: study of people, expertise and sustainable development*, Londres, Psychology Press, 1995, 216 p.

ANDRO Mathieu, *Bibliothèques numériques et crowdsourcing : Expérimentations autour de Numalire, projet de numérisation à la demande par crowdfunding*, 2017, 461 p., Thèse de doctorat en sciences de l'information et de la communication, Saint-Denis, Université Paris 8 Vincennes, sous la direction de Imad Saleh,.

ANDRO Mathieu, SALEH Imad, « Bibliothèques numériques et *gamification* : panorama et état de l'art », in *I2D – Information, données & documents* [en ligne], Volume 53, 2015/4, p. 70-79, consulté le 03 mars 2019, URL : <https://www.cairn.info/revue-i2d-information-donnees-et-documents-2015-4-page-70.htm>.

BOUYÉ Édouard, « Le Web collaboratif dans les services d'archives publics : un pari sur l'intelligence et la motivation des publics », in *Gazette des archives*, [en ligne], n° 227, année 2012, p. 125-136, consulté le 03 mars 2019, URL : https://www.persee.fr/doc/gazar_0016-5522_2012_num_227_3_4974.

BROUDOUX Évelyne, « Indexation collaborative: traces de lecture et constitution de communautés », in *Bibliothèques 2.0 à l'heure des medias sociaux*, Editions du Cercle de la librairie, 2012, p.125-134.

BURGER-HELMCHEN Thierry, PENIN Julien, « Crowdsourcing : définition, enjeux, typologie », in *Management & Avenir*, [en ligne], n°41, 2011/1, p. 254-269, consulté le 28 février 2019, URL : <https://www.cairn.info/revue-management-et-avenir-2011-1-page-254.htm>.

CARLETTI Laura, McAULEY Derek, PRICE Dominic, [et al.], « Digital Humanities and Crowdsourcing: An Exploration », in *Museums and the Web 2013*, N. Proctor & R. Cherry (eds). Silver Spring, [en ligne], février 2013, consulté le 29 janvier 2019, URL : <https://mw2013.museumsandtheweb.com/paper/digital-humanities-and-crowdsourcing-an-exploration-4/>.

CASEMAJOR LOUSTAU Nathalie, « La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? », in *Études de communication*, 36, 2011, [en ligne], consulté le 31 mars 2019, URL : <http://journals.openedition.org/edc/2532>.

CASILLI Antonio A., CARDON Dominique, *Qu'est-ce que le Digital Labor ?*, INA éditions, 2015, p. 16. , in HERBILLE Gauthier, MAZET Jérémy, PETIT Axel, « Transcribe Bentham : recherche historique et crowdsourcing », in *Innovation pédagogique*, 3 janvier 2017, [en ligne], consulté le 3 mars 2019, URL : <https://www.innovation-pedagogique.fr/article1329.html>.

CAVALIÉ Hélène, « Plaidoyer pour une indexation collaborative », in *L'indexation matière en transition : de la réforme de Rameau à l'indexation automatique*, Étienne Cavalie dir., Éditions du Cercle de la Librairie, Collection Bibliothèques, 2019 (à paraître).

CHUPIN Lisa, « Documentarisation participative et médiation du patrimoine scientifique numérisé. Le cas des herbiers », in *Études de communication*, 46, 2016, [En ligne], consulté le 29 mars 2019, URL : <http://journals.openedition.org/edc/6499>.

CHUPIN Lisa, *Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés*, 2017, 220 p., Thèse de doctorat en Sciences de l'information et de la communication, Paris, CNAM, sous la direction de Manuel Zacklad.

CRISTIANO Agostino, « Museum Crowdsourcing as Playful Labour », in *ICOFOM Study Series*, 43a, 2015, [en ligne], consulté le 29 janvier 2019, URL : <http://journals.openedition.org/iss/545>.

DUN Stuart, HEDGES Mark, *Crowd-Sourcing Scoping Study. Engaging the Crowd with Humanities Research*, Londres, Centre for e-Research, Department of Digital Humanities, King's College London, Novembre 2012, 56 p., [en ligne], consulté le 15 mars 2019, URL : <https://stuardunn.files.wordpress.com/2013/04/crowdsourcing-connected-communities.pdf>.

DURIEUX Valérie, « Collaborative tagging et folksonomies. L'organisation du web par les internautes », in *Les Cahiers du numérique*, vol. 6, n° 1, 2010, p. 69-80.

EIDELMAN Jacqueline (dir.), *Rapport de la mission « Musées du XXI^e siècle »*, Paris, Volume 1, Synthèse, Février 2017, 173 p.

ELLIS Sally, « A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship », in *Libri*, 64(1), 2014, p. 1-10.

FONT Frederic, SERRA Joan, SERRA Xavier, « Folksonomy-based tag recommendation for collaborative tagging systems », in *International Journal on Semantic Web and Information Systems*, n° 9-2, 2013, p. 1-30., [en ligne], consulté le 9 mars 2019, URL : <http://mtg.upf.edu/node/2808>.

FRANCIS Elie, QUESNEL Odile, « Indexation collaborative et folksonomie », in *Documentation-Sciences de l'Information*, Vol. 64, 2007/1, p. 58-63.

GUÉRY Bastien, « Un Inventaire participatif du patrimoine national des orgues », in *Participez / Participer. Patrimoines et dispositifs participatifs*, Journées professionnelles, Paris, 5 et 6 octobre 2017.

HEDGES Mark, DUNN Stuart, *Academic crowdsourcing in the Digital Humanities : Crowds, Communities and Co-production*, Londres, Chandos Publishing, Chandos Information Professional Series, 2017, 190 p.

HOOLAND Van Seth, MENDEZ RODRIGUEZ Eva & BOYDENS Isabelle, « Between Commodification and Engagement: On the Double-Edged Impact of User-Generated Metadata within the Cultural Heritage Sector », in *Library Trends*, 59 (4), 2011, p. 707–720.

KOHLE Hubertus, « Artigo : social image tagging pour les œuvres d’art », in Joyeux-Prunel Béatrice (Hg.), *L’art et la mesure. Histoire de l’art et méthodes quantitatives*, Paris, 2010, p. 153-162.

LABIA Camille, *Les humanités numériques au service d’un projet de recherche collaborative en archéologie*, septembre 2018, 70 p., Mémoire de stage, Paris, École du Louvre, sous la direction de Françoise Dalex et Christophe Leclercq.

Laboratoire Société Numérique, « Indexation collaborative : quand les institutions culturelles font appel au public pour annoter les œuvres, les documents et les archives », in *Mission Société Numérique*, Action Publique, Dossiers, Pratiques, Territoires, n°3, septembre 2018, [en ligne], consulté le 9 janvier 2019, URL : <https://labo.societenumerique.gouv.fr/2018/09/03/indexation-collaborative-insitutions-culturelles-appel-public-annoter-oeuvres-documents-archives/>.

LEBRETON Claire, *Bibliothèques, tags et folksonomies. L’indexation des bibliothèques à l’ère sociale*, mars 2008, 117 p., Mémoire d’étude, Paris, Département des ressources documentaires de l’ENSSIB, sous la direction de Nicolas MORIN.

LECOMTE Béatrice, *Europeana comme modèle de portail de diffusion des archives et du patrimoine*, mai 2015, 24 p., EBSI, Université de Montréal, dans le cadre du cours SCI6850 – Recherche individuelle, présenté à Monsieur Yvon Lemay.

LÉVY François, NAZARENKO Adeline, GUISSÉ Abdoulaye, « Annotation, indexation, et parcours de documents numériques. Consulter, explorer et maintenir les textes réglementaires », in *Documents numériques*, Vol. 13, 2010, p. 121-152, [en ligne], consulté le 03 mars 2019, URL : <https://www.cairn.info/revue-document-numerique-2010-3-page-121.htm>.

LÉVY Pierre, *L’Intelligence collective. Pour une anthropologie du cyberspace*, Paris : La Découverte, 1994, 245 p.

LIMON-BONNET Marie-Françoise, BOUYÉ Édouard, « Indexation et transcription collaboratives dans le monde en réseau des archives aujourd’hui : réussites et enjeux d’avenir », in *Participez / Participer. Patrimoines et dispositifs participatifs*, Journées professionnelles, Paris, 5 et 6 octobre 2017.

MILLOT Glen, NEUBAUER Claudia, STORUP Bérangère. *La recherche participative comme mode de production de savoirs. Un état des lieux des pratiques en France*, Une étude menée par la Fondation Sciences Citoyennes avec le soutien de la Fondation de France, 2012-2013, 94 p., [en ligne], consulté le 12 décembre 2018, URL : https://sciencescitoyennes.org/wp-content/uploads/2013/05/FSC-recherche_participative-FdF.pdf.

MOIREZ Pauline, MOREUX Jean Philippe, JOSSE Isabelle, *État de l'art en matière de crowdsourcing dans les bibliothèques numériques*, Livrable L-4.3.1 du projet de R&D du FUI 12 pour la conception d'une plateforme collaborative de correction et d'enrichissement des documents numérisés, février 2013, 77 p., [en ligne], consulté le 03 février 2019, URL : http://www.bnf.fr/documents/crowdsourcing_rapport.pdf.

NEROULDIS Ariane, *Le Crowdsourcing appliqué aux archives numériques : concepts, pratiques et enjeux*, Juin 2015, 109 p., Mémoire du Master Archives numériques, Lyon, Université de Lyon, sous la direction de Céline GUYON.

NOIRET Serge, « La Digital History: histoire et mémoire à la portée de tous », in *Ricerche Storiche*, 1, vol. 41, January-April. 2011, p. 111-148.

NOV Oded, NAANAM Nor, & YE Chen, « Motivational, structural and tenure factors that impact online community photo sharing », in *Proceedings of the Third International ICWSM Conference*, San Jose, California (USA), 17-20 May 2009.

OOMEN Johan, AROYO Lora, « Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges », in *Proceedings of the Fifth International Conference on Communities and Technologies*, Brisbane, QLD, Australia, June 29 - July 2, 2011, 12 p., [en ligne], consulté le 03 février 2019, URL : http://www.iisi.de/fileadmin/IISI/upload/2011/p138_oomen.pdf.

PECCATTE Patrick, « Une plate-forme sociale pour la redocumentarisation d'un fonds iconographique », in *Traitements et pratiques documentaires : vers un changement de paradigme ?*, Actes de la deuxième conférence « Document numérique et Société », sous la direction d'Evelyne Broudoux et Ghislaine Chartron, ADBS, Collection : Sciences et techniques de l'information, 2008, 11 p.

PÉDAUQUE Roger T. (collectif), *La redocumentarisation du monde*, Toulouse, Cepaduès éditions, 2007, 213 p.

PRIME-CLAVERIE Camille, MAHE Annaïg, *Le défi de l'interopérabilité entre plates-formes pour la construction de savoirs augmentés en sciences humaines et sociales*, ISTE éditions. Ecrilecture augmentée dans les communautés scientifiques, 2017, 17 p. [en ligne], consulté le 19 février 2019, URL : https://iste-editions.fr/.sic_01511618.

PROULX Serge, GARCIA José Luis, HEATON Lorna, *La contribution en ligne : pratiques participatives à l'ère du capitalisme informationnel*, Québec, Presses de l'Université du Québec, 2014, 242 p.

PUIG Vincent, PUIG Vincent, « Pratiques contributives à l'âge des données », in CARAYOL Valérie, MORANDI Franc (éd.), *Le tournant numériques des sciences humaines et sociales*, Publication de la Maison des Sciences de l'Homme d'Aquitaine, 2015, 129 p.

RIDGE Mia, « Playing with Difficult Objects – Game Designs to Improve Museum Collections », in *Museums and the Web 2011: Proceedings*, J. Trant and D. Bearman (eds), Toronto, Archives & Museum Informatics, 31 mars 2011, 88 p., [en ligne], consulté le 19 février 2019, URL :

https://www.museumsandtheweb.com/mw2011/papers/playing_with_difficult_objects_game_designs_to.

SALAÜN Jean-Michel, « La redocumentarisation, un défi pour les sciences de l'information », in *Études de communication*, n° 30, 2007, p. 13-23.

SIMON Nina, *The participatory museum*, Santa Cruz, California: Museum 2.0, 2010, 352 p., [en ligne], consulté le 4 février 2019, URL : <http://www.participatorymuseum.org/read/>.

SIMON Rainer, BARKER Elton, ISAKSEN Leif and DE SOTO Cañamares Pau, “Linking early geospatial documents, one place at a time: annotation of geographic documents with Recogito”, in *e-Perimtron*, 10(2), 2015, pp. 49-59.

SMITH-YOSHIMURA Karen, *Social Metadata for Libraries, Archives, and Museums: Executive Summary*, Dublin, Ohio: OCLC, 2012, 20 p., [en ligne], consulté le 14 janvier 2019, URL : <http://www.oclc.org/content/dam/research/publications/library/2012/2012-02.pdf>.

SUROWIECKI James, *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economics, society and nations*, New York, Doubleday, 2004, 320 p.

TRELEANI Matteo, *La recontextualisation des archives audiovisuelles*, Montréal, Les Presses de l'Université de Montréal, coll. Parcours Numériques, 2014, 210 p., [en ligne]. (Consulté le 4 février 2019, URL : <http://www.parcoursnumeriques-pum.ca/la-recontextualisation-des-archives-audiovisuelles>).

VIKHROVA Anne, *L'évaluation de la méthode du crowdsourcing pour la transcription de manuscrits*, 2017, 331 p., Thèse, Laboratoire de Linguistique et Didactique des Langues Étrangères et Maternelles, sous la direction de Thomas Lebarbe.

WYMAN Bruce, CHUN Susan, CHERRY Rich [et al.], « Steve.museum: An Ongoing Experiment in Social Tagging, Folksonomy, and Museums », in J. Trant and D. Bearman (eds.), *Museums and the Web 2006: Proceedings*, Toronto: Archives & Museum Informatics, mars 2006, [en ligne], consulté le 4 janvier 2019, URL: <http://www.archimuse.com/mw2006/papers/wyman/wyman.html>.

ZACKLAD Manuel, CHUPIN Lisa, « Le crowdsourcing scientifique et patrimonial à la croisée de modèles de coordination et de coopération hétérogènes: le cas des herbiers numérisés », in *Canadian Review of Information and Library Science*, Toronto, University of Toronto Press, 39 (3/4), 2015, p.308-328, [en ligne], consulté le 04 avril 2019, URL : <https://muse.jhu.edu/login?auth=0>.

ZORICH Diane, GUNTER Waibel & ERWAY Ricky, *Beyond the Silos of the LAMs: Collaboration Among Libraries, Archives, and Museums*, Dublin, Ohio: Online Computer Library Center, 2008, 59 p.

Billets de blog / presse en ligne

Archives nationales participatives. Projets collaboratifs de transcriptions, annotations et indexations, Carnet Hypothèses.org recensant les projets collaboratifs menés aux Archives Nationales : <https://archivnat.hypotheses.org/>.

CHIGNARD Simon, « Mais pourquoi contribue-t-on ? », in DATA, Données ouvertes, publié le 22 novembre 2012, [en ligne], consulté le 19 janvier 2019, URL : <https://donneesouvertes.info/2012/11/22/mais-pourquoi-contribue-t-on/>.

DANIEL Johanna, « Digital Mellini : Expérimenter un ‘scholars’ workspace’ », in *Isidore et Ganesh*, publié le 20 août 2014, [en ligne], consulté le 19 janvier 2019, URL : <http://johannadaniel.fr/isidoreganesh/2014/08/digital-mellini/>.

HOWE Jeff, « The Rise of Crowdsourcing », in *Wired*, 14, (6), 2006, [en ligne], consulté le 29 janvier 2019, URL : <http://www.wired.com/wired/archive/14.06/crowds.html>.

LE GUERN Isabelle, « Herbiers 2.0. où la Belle au Bois Dormant n’était pas indexée! », in *Infonum Bordeaux Montaigne, Animer et valoriser l’information numérique*, publié le 6 janvier 2016, [en ligne], consulté le 19 janvier 2019, URL : <http://www.isabelleleguern.com/2016/05/30/herbiers-2-0-ou-la-belle-au-bois-dormant-netait-pas-indexee/>.

MOIREZ Pauline, « Bibliothèques, catalogues et crowdsourcing », in *Archives masala*, 20.10.2014 (extrait du cours sur Les évolutions des catalogues, ENSSIB, 6.10.2014), [en ligne], consulté le 09 janvier 2019, URL : <https://archivesmasala.wordpress.com/2014/10/20/bibliotheques-catalogues-et-crowdsourcing/>.

O’REILLY Tim, BATTELLE John, « What is web 2.0? Design Patterns and Business Models for the Next Generation of Software », in *O’Reilly*, publié le 9 septembre 2005, [en ligne], consulté le 19 avril 2019, URL : <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>.

PINEAU Karl, « CaptchA-N : Vers un système de validation de l’indexation collaborative », in *Données culturelles. Carnet de recherche sur les données et ontologies patrimoniales*, publié le 11 décembre 2018, [en ligne], consulté le 14 février 2019, URL : <https://dculturelles.hypotheses.org/33>.

RIDGE Mia, « Notes from ‘Crowdsourcing in the Arts and Humanities’ », in *Open Objects*, publié le 17 avril 2013, [en ligne], consulté le 24 février 2019, URL : <http://openobjects.blogspot.co.uk/2013/04/crowdsourcing-in-arts-and-humanities.html>.

RIDGE Mia, « Frequently Asked Questions about crowdsourcing in cultural heritage », in *Open Objects*, publié le 3 juin 2012, [en ligne], consulté le 24 février 2019, URL : <http://www.openobjects.org.uk/2012/06/frequently-asked-questions-about-crowdsourcing-in-cultural-heritage/>.

TEXIER Bruno, « Projet PhotosNormandie : dix ans d’indexation collaborative », in *Archives patrimoine*, publié le 24 janvier 2017, [en ligne], consulté le 19 janvier 2019, URL : <https://www.archimag.com/archives-patrimoine/2017/01/24/projet-photos-normandie-dix-ans-indexation-collaborative>.

Projets de crowdsourcing mobilisés lors de notre recherche

Testaments de Poilus lancé en 2018 par les Archives nationales et les Archives départementales des Yvelines.

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

RECITAL (Transcription participative des REgistres de la Comédie-ITALienne de Paris au XVIII^e siècle) lancé en 2018 par l'université de Nantes.

URL : <http://recital.univ-nantes.fr/#/>.

Inventaire des orgues lancé en 2017 par le ministère de la Culture et Fédération Francophone des Amis de l'Orgue.

URL : <https://www.inventaire-des-orgues.fr>.

Europeana 1914-1918 lancé du 9 au 16 novembre 2016 par les Archives de France, la Bibliothèque nationale de France et Europeana.

URL : www.europeana1914-1918.eu/fr.

Recogito développé en 2016 par Pelagios Commons sous la direction l'Australian Institute of Technology, l'Université d'Exeter, l'Open University et la Fondation Andrew W. Mellon.

URL : <https://recogito.pelagios.org/>.

Morts pour la France de la Première Guerre mondiale lancé en avril 2015 par la Direction des Patrimoines, de la Mémoire et des Archives.

URL : <http://www.memoiredeshommes.sga.defense.gouv.fr/fr/>.

Digital Muret lancé en 2014 par l'INHA et la Bibliothèque nationale de France.

URL : <https://www.zooniverse.org/projects/inha/digital-muret>.

Transcrire. Site de transcription collaborative en ethnologie (carnets manuscrits de terrain) lancé en 2013 par le Consortium Archives des ethnologues.

URL : <http://transcrire.huma-num.fr/>.

Rijksstudio a été créé en 2013 par le Rijksmuseum pour permettre la découverte des collections du musée en même temps que celui-ci rouvrirait ses ports après dix ans de fermeture.

URL : <https://www.rijksmuseum.nl/en/rijksstudio>.

Les Herbonautes lancé en septembre 2012 par le Muséum national d'Histoire naturelle de Paris, RECOLNAT et TelaBotanica.

URL : <http://lesherbonautes.mnhn.fr>.

ARTigo lancé en 2012 par le Groupe IT des Sciences Humaines, l'Institut d'informatique, l'Institut d'Histoire de l'Art, l'Institut de Philologie et la LMU de Munich.

URL : <https://www.artigo.org/about.html>.

CORRECT est une plateforme de correction d'OCR mise en service en a 2012 et fermée en 2015 par la Bibliothèque nationale de France.

URL : site fermé.

Old Weather est un projet de transcription collaborative développé sur la plateforme Zooniverse en Octobre 2010 par le National Maritime Museum, les Archives nationales de Grande-Bretagne, l'Université d'Oxford, l'Oxford Martin School l'ACRE (International Atmospheric Circulation Reconstructions over the Earth).

URL : <https://www.oldweather.org/>.

Transcribe Bentham lancé en avril 2010 par l'University College of London (UCL).

URL : http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham.es.

Digital Mellini lancée 2010 par le Getty Research Institute.

URL : <http://www.getty.edu/research/mellini/>.

Norfolk Medieval Graffiti Survey est un projet d'archéologie communautaire créé en 2010 et achevé en juin 2012 porté par une association du même nom.

URL : <http://www.medieval-graffiti.co.uk/index.html>.

Trove lancé en 2008 par la Bibliothèque nationale d'Australie.

URL : <http://trove.nla.gov.au/>.

Scottish Words and Place Names (SWAP) a été lancé en 2008, clos en 2011, par la School of Critical Studies du College of Arts de l'Université de Glasgow, le National e-Science Center (NeSC) et le Scottish Language Dictionary (SLD).

URL : <https://swap.nesc.gla.ac.uk/>.

Flickr: The Commons lancé en juin 2008 (projet pilote).

URL : <https://www.flickr.com/>.

Bibliothèque du Congrès –Flickr est une expérimentation initiée en 2007 visant à permettre l'identification de fonds photographiques issus de cette institution.

URL : https://www.flickr.com/photos/library_of_congress/.

Galaxy Zoo.org lancé en 2007 par International University Collaboration.

URL : <http://zoo1.galaxyzoo.org>.

PhotosNormandie a été initié en janvier 2007 pour améliorer les légendes de 3000 photographies et de 300 films relatifs à la bataille de Normandie par Patrick Pecatte et Michel Le Querrec.

URL : <https://www.flickr.com/photos/photosnormandie/>.

Biodiversity Heritage Library lancé en 2006 par un consortium de douze bibliothèques américaines et britanniques possédant de riches fonds d'histoire naturelle afin de créer un portail d'accès à toutes leurs ressources.

URL : <https://www.biodiversitylibrary.org/>.

Steve Museum lancé en 2005 par 12 institutions culturelles dont le Guggenheim Museum, ou le Metropolitan Museum of Art.

URL : <https://sourceforge.net/projects/steve-museum/>.

Ancient Lives lancé en 2001 et coordonné par l'Université d'Oxford, en collaboration avec des chercheurs spécialisés en papyrologie.

URL : www.papyrology.ox.ac.uk/Ancient_Lives/.

Wikipédia lancé en 2001 par la Fondation Wikipédia.

URL : <http://www.wikipedia.org/>.

Garden Birdwatch est un projet de sciences citoyennes mis en œuvre en 1994 par le British Trust of Ornithology (BTO) qui s'est renouvelé avec le Web 2.0.

URL : <https://www.bto.org/our-science/projects/gbw/about/background>.

