

**Pratiques participatives et patrimoine numérisé :
le cas des manuscrits
Réflexion en vue d'un projet de transcription collaborative
en Ville de Genève**

mirabilia

**Travail de Master réalisé en vue de l'obtention du
Master HES en Sciences de l'Information**

par :

Florence BURG Y

Sous la direction de :

Françoise DUBOSSON, Chargée de Cours HES

Genève, le 17 août 2020

**Filière Information Documentaire
Haute école de gestion de Genève (HEG-GE)**

Remerciements

De nombreuses personnes ont contribué à la réalisation de ce travail.

Je souhaiterais tout d'abord remercier Mme Françoise Dubosson et M. Alexandre Boder pour m'avoir soufflé l'idée de sujet et m'avoir accompagnée tout au long de mes recherches.

Ensuite, un grand merci à Mme Véronique Lombard, à M. Olivier Deslarzes et aux membres du Comité éditorial de la plateforme Mirabilia, pour leur confiance et leur collaboration.

Je souhaite également remercier les personnes qui m'ont fait part de leur expertise, M. David Glaser, Mme Barbara Roth, Mme Liliane Regamey, Mme Béatrice Joyeux-Prunel, M. Simon Gabay, Mme Alexa Renggli, M. Ben W. Brumfield, Mme Sara Brumfield, Mme Paule Hochuli, Mme Dominique Chatillon, Mme Noëlle Corboz, M. Alexis Rivier et Mme Sylvie Treglia-Détraz.

Un immense merci encore à mes chers « usagers potentiels », dont je préserve l'anonymat, mais qui se reconnaîtront, pour m'avoir accordé du temps et de l'attention en cette période de crise. Sans eux et leurs multiples contributions, ce travail n'aurait pu voir le jour.

Enfin, un grand merci à mon père, relecteur de haut vol, et à ma famille, mes amis, mes colocataires et mon partenaire pour leur soutien au quotidien.

Résumé

Ce travail s'inscrit dans une réflexion générale autour de la plateforme de valorisation du patrimoine genevois numérisé Mirabilia, et de sa dimension participative. L'objectif principal est de proposer une étude concernant la valorisation participative des documents manuscrits et des pistes de solutions visant à construire un projet allant dans ce sens, en prenant en compte les spécificités du contexte genevois, de ses institutions et de ses publics. Il s'agit d'une part de comprendre les besoins et attentes des tenants du projet et, d'autre part, d'aller à la rencontre d'utilisateurs potentiels pour déterminer comment un tel projet peut les intéresser et leur être utile.

La méthodologie utilisée s'inspire du *Design Thinking* et du *User Experience*. Elle se base sur des recherches personnelles (lectures, observations participatives et entretiens avec des experts), sur une analyse des tenants du projet (questionnaires), des publics (*persona* et questionnaires), et sur une intégration des utilisateurs à la construction du projet (tri de cartes, prototypes et tests utilisateurs). Divers outils ont été réalisés au cours de ce projet, tels un guide décisionnel pour la sélection de fonds, une cartographie des parties prenantes, des parcours utilisateurs, une maquette et une marche à suivre. Les circonstances n'ont pas permis la réalisation d'un prototype fonctionnel, mais il constituera une étape cruciale pour la suite du projet.

Ces étapes ont permis de formuler des lignes directrices et des recommandations, et de relever l'importance de la prise en compte des utilisateurs et de leurs opinions, qui ne coïncident pas toujours avec la vision des tenants du projet. Selon leurs profils, les utilisateurs n'ont d'ailleurs pas les mêmes besoins et attentes.

Les « amateurs », qui fréquenteraient la plateforme pour le loisir, rechercheraient une forme de communauté, une possibilité de se rencontrer en groupes d'intérêt et de collaborer à un projet de grande envergure. Ils peuvent s'intéresser à des thématiques très variées, proposées par les institutions, mais ils ont besoin de se sentir légitimes dans leur participation et de voir que leurs contributions sont utiles.

Dans les domaines de l'éducation et de la recherche, il est impératif d'aller chercher le dialogue avec les personnes concernées pour déterminer des thématiques et des fonds à valoriser. Leur intérêt pour la plateforme dépend de leurs sujets de recherche et d'enseignement, et s'ils ne sont pas consultés, ils risquent de ne pas voir d'utilité immédiate à la plateforme.

Il faut donc que les tenants du projet soient prêts à intégrer les utilisateurs dans la construction de ce projet participatif, et à leur laisser de la place, sans pour autant craindre pour l'image de leur institution. Il faut pour cela prendre le temps de définir des objectifs précis pour éviter l'écueil d'un projet « prétexte », sans autre but que de suivre la tendance à la participation. Il faudra ensuite continuer le processus entamé en créant des prototypes d'outils et en les faisant tester à des utilisateurs potentiels, afin de s'assurer qu'ils répondent à leurs besoins et attentes.

Mots-clés : participation – patrimoine – manuscrits – valorisation du patrimoine numérisé – transcription collaborative – crowdsourcing – Ville de Genève – Mirabilia – Design Thinking

Table des matières

Remerciements.....	i
Résumé	ii
Liste des tableaux	vi
Liste des figures.....	vii
Liste des abréviations.....	viii
1. Introduction.....	1
1.1 Cadre général : le virus, le web, et le patrimoine	1
1.2 Contexte du projet.....	1
1.3 Problématique et objectifs	3
2. Méthodologie	4
2.1 <i>Design Thinking</i> et UX.....	4
2.2 Ajustements méthodologiques.....	6
2.2.1 La théorie et la pratique	6
2.2.2 La crise sanitaire et ses incidences	6
2.3 Méthodes choisies.....	7
2.3.1 Lectures et observations participantes.....	7
2.3.2 Entretiens avec des experts.....	7
2.3.3 Questionnaires et tests – DCTN et Comité éditorial	7
2.3.4 <i>Persona</i> et recrutement d’usagers potentiels	8
2.3.5 Questionnaires et tests – usagers potentiels	9
2.3.6 Tri de cartes.....	9
2.3.7 Parcours utilisateurs	9
2.3.8 Prototype et outils.....	9
2.3.9 Marche à suivre	9
3. Revue de la littérature	10
3.1 Pratiques participatives et patrimoine numérisé.....	10
3.1.1 Définitions et typologies.....	10
3.1.2 Recommandations et bonnes pratiques.....	14
3.1.3 Publics et motivations	16
3.1.4 Enjeux	17
3.1.5 La situation helvétique	17
3.2 Le cas des manuscrits	18
3.2.1 Le paléographe et la machine.....	18
3.2.2 Une solution collective : la transcription collaborative	18
3.2.3 Projets existants	20
4. Résultats et analyse	23
4.1 Observations participantes.....	23
4.1.1 Annotated Books Online	23

4.1.2	Anti-Slavery Manuscripts	24
4.1.3	DIY History	24
4.1.4	E-manuscripta	25
4.1.5	Europeana – Transcribathon	26
4.1.6	Harvard Library – Archives of the Gray Herbarium.....	26
4.1.7	Library of Congress : By the people	27
4.1.8	TACT	28
4.1.9	Transcribe Bentham	29
4.1.10	Transcrire.....	29
4.1.11	Éléments à retenir des observations participantes	30
4.2	Entretiens avec des experts	31
4.2.1	David Glaser – responsable de notrehistoire.ch – 16.04.2020	31
4.2.2	Barbara Roth – conservatrice responsable des collections spéciales de la BGE récemment retraitée – 04.05.2020	32
4.2.3	Liliane Regamey – responsable de e-newspaperarchives.ch – 25.05.2020	32
4.2.4	Alexa Renggli – coordinatrice de la plateforme e-manuscripta.ch et responsable de projet pour l’outil de transcription – 04.06.2020.....	33
4.2.5	Éléments à retenir des entretiens	34
4.3	Questionnaires et tests – DCTN et Comité éditorial.....	35
4.3.1	Objectifs, attentes et besoins	35
4.3.2	Publics, usagers potentiels	36
4.3.3	Un projet participatif.....	38
4.3.4	Sélection des documents.....	38
4.3.5	Plateformes testées	40
4.4	Persona et recrutement d’usagers potentiels	40
4.5	Questionnaires et tests – usagers potentiels	42
4.5.1	Enseignement et recherche en Histoire	42
4.5.2	Amateurs	44
4.5.3	Généalogie	47
4.5.4	Enseignement et recherche en Sciences naturelles.....	47
4.5.5	Récapitulation générale	49
4.6	Tri de cartes	50
4.7	Parcours utilisateurs	53
4.8	Prototype et outils	55
4.8.1	Maquette	55
4.8.2	Outils envisageables.....	59
4.8.3	Informations récentes et proposition de solutions	61
4.9	Marche à suivre	63
5.	Recommandations.....	65
5.1	Recommandations techniques	65
5.1.1	Qualité et interopérabilité.....	65
5.1.2	Cadre juridique	65

5.1.3	Communication et collaboration.....	65
5.1.4	« Construire pour » et « construire avec »	65
5.2	Recommandations stratégiques.....	66
5.2.1	Le parapluie et les baleines	66
5.2.2	La vitrine et la porte d'entrée	66
5.2.3	Une communauté plutôt qu'une foule	66
5.2.4	Ouvrir la porte... ..	66
5.2.5	...et sortir de la pièce.....	67
6.	Conclusion	68
	Bibliographie	69
Annexe 1 :	Grille d'expérience	75
Annexe 2 :	Guides d'entretien.....	77
Annexe 3 :	Questionnaire adressé aux responsables de la plateforme Mirabilia	80
Annexe 4 :	Questionnaire adressé au Comité éditorial de la plateforme Mirabilia	87
Annexe 5 :	Persona	94
Annexe 6 :	Questionnaire adressé aux usagers potentiels – Enseignement et recherche en Histoire.....	99
Annexe 7 :	Questionnaire adressé aux usagers potentiels – Amateurs	102
Annexe 8 :	Questionnaire adressé aux usagers potentiels – Généalogistes	105
Annexe 9 :	Questionnaire adressé aux usagers potentiels – Sciences naturelles	108
Annexe 10 :	Prototype papier.....	111
Annexe 11 :	Maquette	112
Annexe 12 :	Maquette modifiée.....	116
Annexe 13 :	Outils de transcription collaborative – Google Sheet.....	120

Liste des tableaux

Tableau 1 : Publics potentiels identifiés par les tenants du projet	36
Tableau 2 : Comparatif d'outils de transcription collaborative.....	60

Liste des figures

Figure 1 : Le processus de <i>Design Thinking</i>	4
Figure 2 : Les huit commandements de l'UX en bibliothèque selon Schmidt et Etches	5
Figure 3 : « Je suis content que nous soyons tous d'accord ! », d'après Jeff Patton	6
Figure 4 : Échelle de la participation citoyenne selon Sherry Arnstein (1969)	11
Figure 5 : Continuum de la participation du public, du projet WTL (2008)	11
Figure 6 : Checklist de Rose Holley	14
Figure 7 : Bénéfices du <i>crowdsourcing</i> pour des institutions culturelles, d'après l'article de Rose Holley (Holley 2010)	15
Figure 8 : Recommandations de l'OCLC pour des projets de <i>crowdsourcing</i>	15
Figure 9 : Typologie des dynamiques de motivations pour le <i>crowdsourcing</i>	17
Figure 10 : <i>Workflow</i> d'un projet de transcription collaborative.....	19
Figure 11 : Capture d'écran de Annotated Books Online	23
Figure 12 : Capture d'écran de Anti-Slavery Manuscripts	24
Figure 13 : Capture d'écran de DIY History.....	24
Figure 14 : Capture d'écran de e-manuscripta	25
Figure 15 : Capture d'écran d'Europeana - Transcribathon.....	26
Figure 16 : Capture d'écran du projet Archives of the Gray Herbarium	26
Figure 17 : Capture d'écran de Library of Congress : By the People.....	27
Figure 18 : Capture d'écran de TACT	28
Figure 19 : Capture d'écran de Transcribe Bentham.....	29
Figure 20 : Capture d'écran de Transcrire.....	29
Figure 21 : Éléments retenus des observations participantes	30
Figure 22 : Éléments à retenir des entretiens avec les experts	34
Figure 23 : Parties prenantes à prendre en compte dans un projet de <i>crowdsourcing</i> en Ville de Genève. Pour chaque projet, les parties prenantes doivent être précisées selon le(s) domaine(s) concerné(s) (par exemple « étudiants en Biologie », « chercheurs en Histoire de l'art » etc.).....	37
Figure 24 : Guide décisionnel pour la sélection de fonds	39
Figure 25 : Exemples de <i>persona</i>	41
Figure 26 : Éléments à retenir pour la future plateforme – mis à jour	49
Figure 27 : Résultats du tri de carte en ligne	52
Figure 28 : Parcours utilisateurs	54
Figure 29 : Page d'accueil de la maquette	55
Figure 30 : Nouvelle version de la page d'accueil de la maquette.....	62
Figure 31 : Nouvelle version du profil utilisateur de la maquette	62
Figure 32 : Marche à suivre pour de futurs projets	64

Liste des abréviations

AEG : Archives d'État de Genève

Ariana : Ariana – musée suisse de la céramique et du verre (membre de Mirabilia)

AVG : Archives de la Ville de Genève (membre de Mirabilia)

BGE : Bibliothèque de Genève (membre de Mirabilia)

BM : Bibliothèques municipales de la Ville de Genève (membre de Mirabilia)

BN : Bibliothèque nationale suisse

CGU : Conditions Générales d'Utilisation

CJB : Conservatoire et Jardin Botaniques (membre de Mirabilia)

CMS : *Content-Management System* (système de gestion de contenu)

DCS : Département de la Culture et du Sport (devenu DCTN en 2020)

DCTN : Département de la Culture et de la Transition Numérique (ex-DCS)

DIP : Département de l'Instruction Publique

FMAC : Fonds Municipal d'Art Contemporain (membre de Mirabilia)

GLAM : *Galleries, Libraries, Archives, Museums* (Galleries, Bibliothèques, Archives, Musées)

LoC : *Library of Congress* (Bibliothèque du Congrès, aux États-Unis)

MAH : Musée d'Art et d'Histoire (membre de Mirabilia)

MEG : Musée d'Ethnographie de Genève (membre de Mirabilia)

MHN : Musée d'Histoire Naturelle (membre de Mirabilia)

OCLC : *Online Computer Library Center*

OCR : *Optical Character Recognition* (reconnaissance optique de caractères)

RTS : Radio Télévision Suisse

TEI : *Text-Encoding Initiative* (initiative pour l'encodage du texte)

UNIGE : Université de Genève

UX : *User Experience* (expérience utilisateur)

VdG : Ville de Genève

WYSIWYG : *What You See Is What You Get* (ce que vous voyez est ce que vous obtenez)

1. Introduction

1.1 Cadre général : le virus, le web, et le patrimoine

Depuis l'avènement du web 2.0, au début des années 2000, les internautes ont pris l'habitude d'agir en acteurs¹, et plus seulement en consommateurs, de l'information en ligne. Des plateformes reposant entièrement sur la création de contenu par ses usagers, comme Facebook ou Wikipédia, se sont rapidement popularisées, ouvrant la porte à des usages jusqu'alors inédits du web (O'Reilly 2005). Ces nouvelles pratiques, basées sur l'intelligence collective et la participation, lui ont valu le nom de web participatif.

En Suisse, comme ailleurs, les institutions culturelles ont suivi la tendance en devenant de plus en plus présentes sur le web. Genève n'est pas en reste, et ses institutions ont su accepter et s'approprier la transition numérique, à travers des campagnes de numérisation et de participation à des bibliothèques numériques, en augmentant l'offre de ressources en ligne et en étant actives sur Wikipédia et les réseaux sociaux (Jacquesson, Von Roten 2019).

Une étude de 2014, reprise par l'Office Fédérale de la Culture, a montré que, cette année-là, « 64% de la population suisse a pratiqué une activité culturelle pendant son temps libre » (OFC 2019). Cependant, cette même étude montre que l'obstacle majeur aux activités culturelles est le manque de temps à disposition (OFC 2019). Une disponibilité en ligne de la culture semble donc nécessaire pour en faciliter l'accès, d'autant plus qu'en 2019, 89% des Suisses entre 16 et 74 ans faisaient un usage quotidien d'Internet (OFS 2020).

La crise sanitaire de 2020 et les bouleversements qui s'en sont suivis ont renforcé ce besoin. Confinés ou semi-confinés chez eux, les citoyens de Suisse et d'ailleurs ont cherché à se divertir et s'instruire, créant une forte demande en matière de culture et de patrimoine en ligne. De nombreuses institutions se sont mobilisées pour proposer une offre axée sur la participation, à l'image du *Stay At Home Challenge* relayé par le Rijksmuseum (Rijksmuseum 2020) ou de la chasse aux œufs virtuelle de la Ville de Genève (Ville de Genève 2020a).

Ces pratiques existaient avant la crise, et elles perdureront après elle. Dans quelle mesure, l'avenir nous le dira. Cette situation a cependant montré que la culture et le patrimoine ont leur place sur le web, et que leur valorisation doit passer par une ouverture à des pratiques participatives, qui permettent aux publics de se les approprier et de leur donner vie.

1.2 Contexte du projet

Comme dit précédemment, les institutions culturelles de la Ville de Genève sont présentes sur le web depuis de nombreuses années, avec entre autres des projets de numérisation visant à la fois une meilleure conservation et une plus large diffusion du patrimoine.

Ces initiatives ne sont qu'un début, et la Ville de Genève « souhaite affirmer son rôle dans le domaine de la numérisation et des données, et développer un discours critique autour de ces enjeux éthiques, que ce soit en matière d'accaparement des données, d'impact écologique et d'enjeux sociaux, d'emploi et de formation » (Ville de Genève 2020b). Symbole fort, le Département de la Culture et du Sport (DCS) devenait, le 1^{er} juin 2020, le Département de la

¹ Ce travail utilise systématiquement le masculin générique pour une question de lisibilité, mais il va de soi que les termes concernés englobent tous les genres.

Culture et de la Transition Numérique (DCTN)², sous la direction du magistrat Sami Kanaan. L'un de ses mots d'ordre, « la mise en valeur d'une ville numérique dans une approche inclusive, participative et créative » (Ville de Genève 2020c), démontre une volonté de lier étroitement citoyens, culture et numérique.

En 2016 déjà, un mandat portant sur la valorisation des collections numériques dans les institutions culturelles (Boder et al., 2016) a servi de base de réflexion pour la création de Mirabilia, une « plateforme digitale interactive didactique et innovante » (Ville de Genève 2020d). Mise en ligne en mai 2020, elle est encore en phase de test, et la communication à son sujet n'a pas encore commencé officiellement. Son but est de mettre en valeur le patrimoine numérisé et conservé par les institutions culturelles de la Ville de Genève, à savoir :

- Les Archives de la Ville de Genève (AVG)³
- L'Ariana – musée suisse de la céramique et du verre
- La Bibliothèque de Genève (BGE)
- Les Bibliothèques municipales de la Ville de Genève (BM)
- Les Conservatoire et Jardin Botaniques (CJB)
- Le Fonds Municipal d'Art Contemporain (FMAC)
- Le Musée d'Art et d'Histoire (MAH)
- Le Musée d'Ethnographie de Genève (MEG)
- Le Muséum d'Histoire Naturelle (MHN)

Les collections présentes sur la plateforme sont extrêmement variées, témoignant ainsi de la richesse des fonds conservés par ces institutions. Mirabilia est la première initiative visant à rassembler les fonds de toutes ces institutions en un seul espace numérique. La collaboration entre institutions est un enjeu crucial pour ce projet, d'autant plus du fait de la diversité de leurs missions, collections et publics.

La dimension interactive de la plateforme s'incarne actuellement dans un jeu de mots-clés permettant de lier des objets qui n'auraient *a priori* rien à voir entre eux, et dans la possibilité pour les utilisateurs de créer leurs propres collections en y ajoutant leurs objets préférés. Les responsables du projet Mirabilia souhaitent donner encore plus de place à la participation, notamment à travers des projets de *crowdsourcing*, qui ont un double intérêt. De tels projets permettent en effet d'une part d'inviter les usagers à la découverte des collections en créant des données les concernant (tags, transcription...), et d'autre part d'augmenter la « trouvabilité » des documents et d'en simplifier l'accès grâce aux données produites.

Un second mandat, débuté en 2020 (Boder, Dubosson, 2020), propose des réflexions à ce sujet et des solutions allant dans ce sens pour divers types de collections. En parallèle, ce Travail de Master se concentre sur le cas des documents manuscrits.

² Pour simplifier la lecture, ce travail utilise le sigle DCTN en tout temps.

³ Les AVG sont un cas particulier, car elles ne dépendent pas du DCTN.

1.3 Problématique et objectifs

Le problème que posent les manuscrits est triple : numérisés, ils ne « racontent » rien tant que l'on n'a pas déchiffré leur contenu écrit, contrairement aux documents iconographiques ou aux objets. En outre, contrairement aux imprimés, ce contenu ne peut être transcrit automatiquement par une machine, les logiciels OCR⁴ récents ne donnant que rarement des résultats satisfaisants avec ce type de documents (Terras 2016). Enfin, la transcription manuelle est une tâche laborieuse qui ne peut être aisément ajoutée aux cahiers des charges des spécialistes de l'information, et il n'est pas courant que les chercheurs se prêtant à cet exercice rendent disponible le fruit de leur labeur sur les sites d'institutions culturelles.

Pour pouvoir les valoriser en ligne de manière participative, il faut donc trouver une solution permettant de rendre le contenu des documents accessible, et la transcription collaborative apparaît alors comme la solution idéale. Cette méthode consiste à mettre à disposition des manuscrits numérisés en demandant à des volontaires de les transcrire à l'aide d'un outil *ad hoc*. Ces transcriptions sont ensuite vérifiées, validées et intégrées au système de gestion documentaire de l'institution, permettant ainsi une recherche plein texte dans le document et facilitant le travail des étudiants et des chercheurs, professionnels ou amateurs.

Si le principe de base semble simple, un tel projet doit être en adéquation avec le contexte, les parties prenantes et les enjeux de l'environnement dans lequel il est déployé. Le présent Travail de Master a pour objectif principal de construire un projet adapté aux spécificités du contexte genevois, de ses institutions et de ses publics.

Ce projet s'inscrivant dans le cadre plus large de Mirabilia et de la collaboration entre institutions, une vision commune doit être partagée par les tenants du projet, à savoir les responsables du projet Mirabilia au sein du DCTN, et les représentants des diverses institutions, qui forment le Comité éditorial de la plateforme. Pour assurer l'engagement de toutes les parties, il est essentiel d'assurer un consensus quant à la raison d'être du projet.

Il est de plus tout aussi important que ce consensus coïncide avec les besoins et attentes des publics. Si personne ne transcrit ou n'utilise les transcriptions produites, le projet n'a pas de sens et n'est qu'une perte de temps et d'argent. Il est donc essentiel d'identifier les publics cibles et d'aller à leur rencontre, afin de rendre ce projet utile et attractif pour eux. L'objectif est de bâtir ce projet non pas seulement pour eux, mais avec eux.

Il s'agit donc dans un premier temps de questionner la demande institutionnelle en la confrontant aux points de vue des diverses personnes en charge de la plateforme. Ceci permettra d'établir des lignes directrices acceptées par tous les tenants du projet.

Dans un second temps, il faut identifier les besoins des publics cibles et leur intérêt pour la transcription collaborative en allant à leur rencontre. Ceci permettra de comprendre comment ce projet peut être utile et intéressant pour de futurs usagers, et de trouver un équilibre avec la demande institutionnelle.

Il sera ainsi possible de poser les bases de ce projet, à la fois avec le Comité éditorial, mais aussi, et surtout, avec ses usagers potentiels, dans le but qu'il réponde autant que possible à leurs besoins et attentes.

⁴ Abréviation de « *Optical Character Recognition* », reconnaissance optique de caractères.

2. Méthodologie

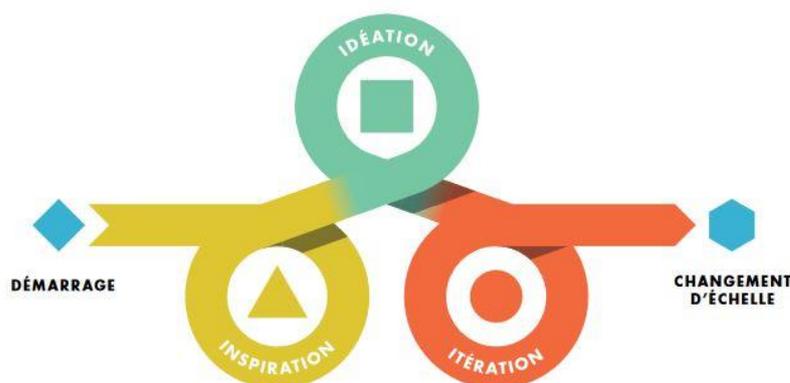
La problématique de ce travail repose avant tout sur des questions humaines plutôt que techniques. Une méthodologie centrée sur les personnes en général et sur les usagers en particulier s'impose alors naturellement, et ce travail s'inspire, dans ce but, des principes du *Design Thinking* et du *User Experience* (UX).

2.1 *Design Thinking* et UX

Démocratisé par l'agence américaine IDEO et la d.school de Stanford, Californie, dans les années 1990 (Biso, Le Naour 2017), le *Design Thinking* est une méthode de travail mais aussi une façon de penser (IDEO 2016). Centré sur l'humain, il implique avant tout d'aller à la rencontre des gens et de leurs besoins, et de développer, tester et améliorer des prototypes qui peuvent y répondre (IDEO 2016). Parmi les valeurs fortes que cette méthode défend, on retient en particulier l'empathie, le fait de remettre l'humain au centre, le fait de « faire », de prototyper, l'itération, donc le droit de se tromper et de réessayer, et enfin la force de l'intelligence collective (Aldana, Dromer Leméni 2019).

De manière générale, on considère que le processus de *Design Thinking* se construit selon trois grandes étapes : l'inspiration, l'idéation et l'itération.

Figure 1 : Le processus de *Design Thinking*



(IDEO 2016, p. 8)

Nicolas Beudon, une référence francophone en matière de *Design Thinking* pour les bibliothèques, définit ces trois étapes ainsi :

1. « Une phase de définition des besoins des utilisateurs finaux à travers l'observation et l'immersion ;
2. Une phase de brainstorming où l'on cherche à produire un maximum d'idées avant de sélectionner les meilleures pour réaliser des prototypes ;
3. Une phase de test et d'amélioration des prototypes avant d'implémenter une solution. »

Il ne s'agit donc pas d'une méthode linéaire, mais bien d'un processus itératif, lors duquel on avance à tâtons (Beudon 2015a), au fil des tests et des retours d'expérience des utilisateurs.

La notion d'« expérience utilisateur » est fondamentale. Le sigle UX, pour *User Experience* a été proposé par Donald Norman, dans les années 1980, et Nicolas Beudon le définit comme « l'ensemble des points de contacts entre un utilisateur et un service, et le ressenti qu'ils impliquent » (Beudon 2017a).

L'UX n'est donc pas qu'une affaire d'informatique ou d'industrie, et cette méthode a également un rôle à jouer pour la bibliothéconomie, comme le montre un ouvrage traduit en français par Nathalie Clot (Schmidt, Etches 2016). Selon ses auteurs, les trois piliers de l'UX sont le fait d'être utile, utilisable et désirable. Ils proposent ensuite les huit commandements de l'UX en bibliothèque, que Nicolas Beudon a représentés ainsi sur son blog :

Figure 2 : Les huit commandements de l'UX en bibliothèque selon Schmidt et Etches



(Beudon 2017b)

Le *Design Thinking* et l'UX ont donc pour but d'intégrer les usagers à toutes les étapes du projet en cherchant à comprendre leurs besoins, leurs ressentis et leur façon de penser. Mettre l'humain au centre implique cependant de mettre également les collaborateurs au centre, car ils sont tout autant essentiels au bon déroulé du projet, et il est important de comprendre leurs besoins et attentes (Biso, Le Naour 2017). Pour adopter cette approche, vis-à-vis des usagers comme des collaborateurs, il existe de très nombreuses possibilités d'exercices, d'ateliers, de tests et de jeux, et la littérature sur le sujet est riche.

La méthodologie de ce travail est principalement basée sur le kit pratique *Le design thinking en bibliothèque* (IDEO 2016), sur deux ouvrages francophones de *Design Thinking* (Biso, Le Naour 2017 ; Aldana, Dromer, Leméni 2019) ainsi que sur l'ouvrage mentionné plus haut qui concerne l'UX en bibliothèque (Schmidt, Etches 2016). Les supports d'un enseignement délivré à la Haute école de gestion de Genève par Nicolas Beudon (Beudon 2019a) ainsi qu'un Travail de Master en Sciences de l'Information récent (Sabatini Choquard 2018) ont également permis de mieux aborder ces méthodes et leurs applications.

2.2 Ajustements méthodologiques

2.2.1 La théorie et la pratique

Comme indiqué plus haut, ce travail s'inspire des principes du *Design Thinking* et du UX, mais ne les suit pas toujours scrupuleusement. En effet, comme le reconnaît Nicolas Beudon, il n'est pas toujours évident d'intégrer totalement le *Design Thinking* dans le contexte d'institutions publiques, mais on peut tout de même l'exploiter et s'en inspirer (Beudon 2015b).

Dans le cas présent, contrairement à ce que préconise le *Design Thinking*, le projet part d'une demande institutionnelle, et non d'un besoin usager, mais l'un des objectifs est de questionner cette demande pour identifier les besoins, attentes et objectifs individuels qu'elle peut receler. On constate effectivement qu'il est fréquent pour une équipe de croire qu'un consensus existe alors que la vision de chacun diffère, comme l'illustre si bien l'image ci-dessous :

Figure 3 : « Je suis content que nous soyons tous d'accord ! », d'après Jeff Patton



(Beudon 2019b)

En outre, la plateforme Mirabilia n'ayant pas encore été dévoilée officiellement au public genevois, elle n'a pas encore d'utilisateurs que l'on pourrait contacter dans le cadre de ce projet. Au lieu de partir de rien, ce travail a pour second objectif de déterminer les types de personnes qu'un tel projet pourrait intéresser et de contacter des individus représentant ces « usagers potentiels » afin d'estimer si le projet est pertinent pour eux et, si oui, selon quelles modalités. Il s'agira ensuite d'associer ces personnes aux discussions et décisions tout au long du projet, comme le préconise le *Design Thinking*.

2.2.2 La crise sanitaire et ses incidences

Un élément imprévu a cependant bouleversé une partie de ce travail, qui devait à l'origine s'articuler autour d'une série de rencontres-ateliers regroupant les différentes parties prenantes, dans l'esprit du *Design Thinking*. La crise sanitaire qui a ébranlé le monde en 2020 a rendu cette approche inapplicable du fait de l'impossibilité de se rencontrer en groupe et de la surcharge de travail à laquelle ont dû faire face les personnes concernées.

N'ayant pas trouvé de solutions idéales, testées et approuvées de *Design Thinking* à distance, j'ai fait les choix méthodologiques qui m'ont semblé appropriés. J'ai dû entre autres m'adapter aux disponibilités des uns et des autres, aux moyens technologiques à leur disposition et à l'arbitrage du DCTN qui déterminait ce que je pouvais demander ou non aux représentants des institutions, en fonction de leur charge de travail.

Je n'ai par exemple pas eu l'occasion de rencontrer le Comité éditorial de Mirabilia en personne ou par visioconférence. De ce fait, j'ai dû me baser uniquement sur leurs retours par écrit, ce qui induit un biais. En outre, si j'ai pu rencontrer les usagers potentiels à distance, j'ai dû limiter ces sessions à des discussions brèves. Dans un cas comme dans l'autre, cela implique notamment que je n'ai pas pu organiser de *brainstorming* à proprement parler, et que j'ai alors choisi de mêler les phases d'inspiration et d'idéation.

Les méthodes choisies ainsi que les problématiques liées à la crise sanitaire ont de plus rallongé considérablement le temps nécessaire à la collecte des informations. Souhaitant aller aussi loin que possible dans ce projet, j'ai choisi de prendre les devants sur certaines étapes, comme par exemple le recrutement des usagers potentiels. J'ai effectué des réajustements par la suite, mais il est évident que mes choix induisent un autre biais, notamment du fait que le recrutement a essentiellement eu lieu au sein de mon propre réseau.

Ces ajustements m'ont permis de poser des bases suffisamment solides pour ce projet, malgré les circonstances et les biais possibles.

2.3 Méthodes choisies

2.3.1 Lectures et observations participantes

La première étape de ce travail a été de se renseigner sur les pratiques participatives dans les institutions culturelles, avec un focus sur les manuscrits et la transcription collaborative, afin d'effectuer une [revue de la littérature](#).

Elle a également consisté en des tests de plateformes de transcription collaborative existantes. Ces tests ont été construits comme des observations participantes, c'est-à-dire qu'il a fallu se mettre dans la peau d'une personne désireux de découvrir la transcription collaborative et évaluer son expérience à chaque étape. Une [grille d'expérience](#), inspirée de celle proposée dans l'ouvrage *Design thinking : accélérez vos projets par l'innovation collaborative* (Biso, Le Naour 2017, p. 68) a été élaborée à cet effet.

Cette grille a également servi à d'autres étapes du travail, afin de comparer l'expérience vécue avec celles du Comité éditorial et des usagers potentiels sur une base commune.

2.3.2 Entretiens avec des experts

En parallèle, des experts pouvant s'intéresser à ce projet ont été identifiés et contactés, afin de recueillir leurs points de vue, conseils et retours d'expériences. Ces rencontres ont pris la forme d'entretiens semi-dirigés, pour lesquels des [guides d'entretien](#) ont été réalisés. Ces guides ont cependant rarement été suivis scrupuleusement, car l'on a préféré laisser les discussions suivre leur cours naturel, mais tous les points mentionnés ont été abordés.

2.3.3 Questionnaires et tests – DCTN et Comité éditorial

Afin de capter et comprendre les divers points de vue des personnes concernées, des questionnaires ont été diffusés auprès des responsables de Mirabilia. L'objectif étant de laisser aux répondants la place de s'exprimer, et de remplacer ainsi, autant que faire se peut, les *brainstormings* qui n'ont pu avoir lieu, on a privilégié des questions aussi ouvertes que possible.

Pour les représentants du DCTN, le [questionnaire](#) (annexe 3) couvre les points suivants :

- Les projets participatifs, afin de connaître leur expérience en la matière et, peut-être, leurs aprioris, mais aussi d'estimer les niveaux de participation envisageables pour le projet. Les différents niveaux de participation proposés s'inspirent d'un ouvrage de Raphaëlle Bats, dont il sera fait mention [plus loin](#).
- Les publics visés, afin d'obtenir une liste des types de publics envisagés pour ce projet. Les répondants doivent également choisir l'un de ces publics et remplir une carte d'empathie le concernant, en se mettant dans sa peau. Cet exercice, qui aurait bien entendu été plus intéressant à faire lors d'ateliers en groupe, permet de mieux comprendre comment les répondants perçoivent les besoins et la façon de penser du public visé, et ce qu'ils pensent que le projet peut apporter.
- Les idées de valorisation des transcriptions produites ainsi que quelques questions techniques, afin de comprendre quels points ont déjà été déterminés en amont et de mieux évaluer le champ d'action possible.
- Les objectifs, besoins et attentes concernant le projet, sous forme de tableau à remplir librement, dans le but, d'une part, de comprendre les éléments qui sous-tendent la demande institutionnelle, et d'autre part d'estimer dans quelle mesure un consensus entre les tenants du projet est déjà en place.

Pour les membres du Comité éditorial, le [questionnaire](#) (annexe 4) traite des points suivants :

- Les projets participatifs : idem que pour les représentants du DCTN
- Les publics : idem que pour les représentants du DCTN
- Les documents ou fonds envisageables pour un projet de transcription collaborative, afin d'obtenir un premier aperçu de ce que les institutions peuvent proposer, selon des critères de sélection minimaux.
- Les objectifs, besoins et attentes : idem que pour les représentants du DCTN.

En plus des questionnaires, les personnes contactées ont reçu la [grille d'expérience](#) mentionnée plus haut, afin de se familiariser avec la transcription collaborative en testant une plateforme existante et en notant les fonctionnalités appréciées et les difficultés rencontrées. Pour limiter la charge de travail imposée, ces tests sont facultatifs et les répondants qui le souhaitent peuvent librement choisir une des plateformes proposées, à savoir [Anti-Slavery Manuscripts](#), [e-manuscripta](#), [Europeana](#), [Library of Congress : By the people](#) et [TACT](#). Ces dernières ont été sélectionnées en fonction de la diversité de leurs fonctionnalités, et du fait qu'il s'agit de plateformes récentes à l'apparence engageante.

2.3.4 Persona et recrutement d'usagers potentiels

Des réponses aux questionnaires doit ressortir une liste de types des publics visés par le projet. Cette liste permet d'établir des [persona](#), qui servent de base pour identifier et contacter des usagers potentiels correspondants.

Du fait des circonstances particulières, ce recrutement a eu lieu en amont, et des réajustements ont été effectués par la suite, une fois les réponses reçues et les [persona](#) réalisées.

2.3.5 Questionnaires et tests – usagers potentiels

Les usagers potentiels recrutés ont tous reçu la [grille d'expérience](#) et un questionnaire. Pour eux, le test d'une plateforme est obligatoire, afin qu'ils puissent découvrir ce qu'est la transcription collaborative, noter les éléments de la plateforme testée qui leur plaisent le plus ou le moins, et répondre au questionnaire en connaissance de cause. Le choix de cette plateforme demeure libre, parmi celles proposées (idem que pour les tenants du projet).

Les [questionnaires](#) (annexes 6 à 9) varient légèrement selon le profil des usagers auxquels ils s'adressent, mais tous concernent les problématiques suivantes :

- L'intérêt perçu pour un tel projet à Genève
- Les besoins auxquels ce projet pourrait répondre
- Les attentes des usagers potentiels
- Leurs craintes

Cette démarche permet non seulement de mieux estimer la direction que ce projet doit prendre, mais aussi de comparer les réponses des usagers à celles du Comité éditorial et des représentants du DCTN, afin de trouver un équilibre qui convienne à chacun.

2.3.6 Tri de cartes

Les usagers potentiels ont ensuite été invités à participer à un tri de cartes en ligne, créé sur la base de informations récoltées. Le but de cet exercice est de classer divers éléments liés au projet selon l'utilité et l'intérêt perçus. Ceci permet alors de définir lesquels inclure en priorité, et lesquels exclure.

Les membres du Comité éditorial et les représentants du DCTN n'ont pas participé à ce tri de cartes, mais ils ont été informés des points qui soulevaient des questionnements.

2.3.7 Parcours utilisateurs

Les étapes précédentes ont donné lieu à l'élaboration de parcours utilisateurs, qui permettent de montrer les parcours envisagés pour les différents types de publics ciblés.

2.3.8 Prototype et outils

Tous ces éléments ont ensuite servi à la réalisation d'une maquette de plateforme, que les usagers potentiels ainsi que les membres du Comité éditorial et les représentants du DCTN ont pu tester. Sur la base de leurs retours, elle a été modifiée et adaptée.

Le temps a manqué pour mettre en place un prototype permettant un test en situation réelle. Cette étape demeure cependant cruciale, et ce prototype devra être testé par des usagers potentiels lors de sa réalisation, idéalement à l'aide de la [grille d'expérience](#), afin de le modifier et de l'ajuster par itérations. Ce travail propose une sélection d'outils de transcription collaborative adéquats ainsi que des recommandations allant dans ce sens.

2.3.9 Marche à suivre

Enfin, ces différentes étapes ont permis d'établir une marche à suivre pour l'élaboration de ce projet par ses responsables. Cette marche à suivre se veut généralisable, afin de servir à de futurs projets du même type également.

3. Revue de la littérature

Cette revue de la littérature est construite en deux parties. La première traite des pratiques participatives, avec un focus sur le *crowdsourcing* et sa mise en pratique dans des institutions patrimoniales. La seconde concerne plus spécifiquement les manuscrits et la transcription collaborative, son fonctionnement et les projets existants.

3.1 Pratiques participatives et patrimoine numérisé

En cette période, que certains appellent « âge de la participation », expression découlant à la fois des évolutions politiques de la seconde moitié du XX^{ème} siècle et de l'impact sociétal des nouvelles technologies au début du XXI^{ème} siècle (Black, 2018), les institutions culturelles dans leur ensemble sont amenées à repenser leur rapport aux publics. La notion de participation est dans toutes les bouches, et des initiatives voient le jour à travers le monde, interrogeant sur ce que recouvre réellement ce concept, que nous allons tenter de définir ici.

3.1.1 Définitions et typologies

3.1.1.1 Participation

L'ouvrage de Raphaëlle Bats, *Construire des pratiques participatives en bibliothèque* (Bats 2015), est un excellent point d'entrée pour comprendre ce que l'on entend par « participation ». Selon l'autrice, ce concept « implique un principe d'égalité qui est au cœur des sociétés démocratiques, depuis l'agora grecque. » (Bats 2015, p. 18). Il s'agit donc avant tout de questionner la hiérarchie habituelle entre les décideurs et ceux pour qui l'on décide, la légitimité et la perméabilité de celle-ci. Ce processus trouve sa place dans de nombreux domaines de la vie politique, sociale, et bien sûr culturelle.

On reconnaît généralement trois niveaux de participation : information, délibération et décision. Ces niveaux doivent cependant être nuancés en fonction de quatre autres modalités : temps (récurrence et régularité), acteurs, sens de la mise en œuvre (*bottom-up* ou *top-down*) et dispositifs (assemblées, conseils consultatifs etc.) (Bats 2015).

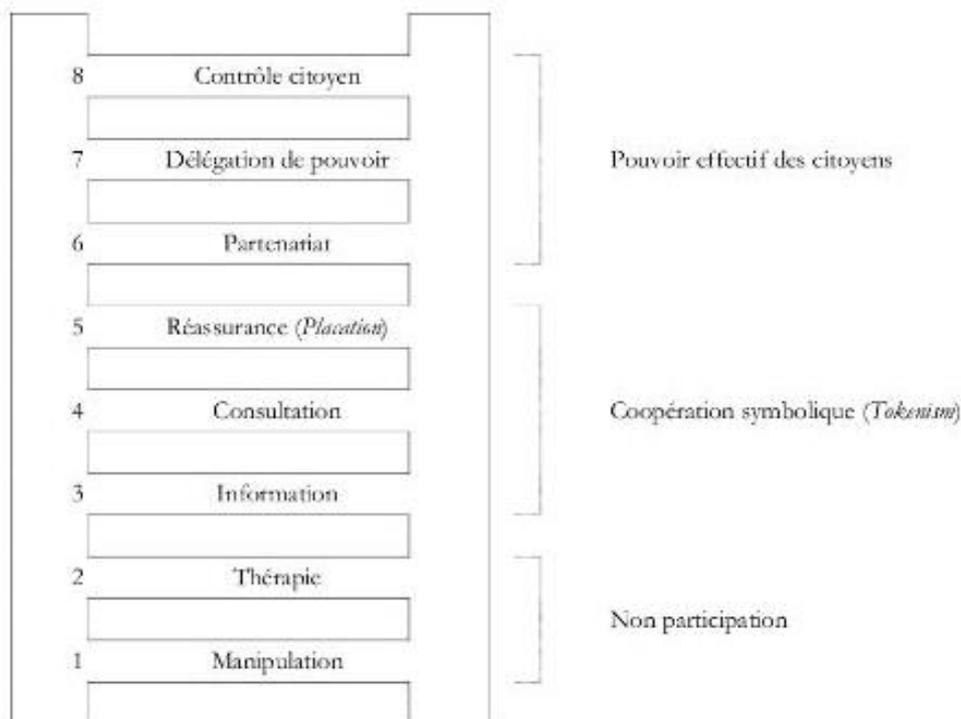
Il est cependant parfois complexe de saisir la nuance entre participatif et collaboratif. Un mémoire de recherche, écrit par Ariane Néroutidis, distingue en effet les projets participatifs, où les participants contribuent à un projet monté et tenu par des décideurs, et les projets collaboratifs, « pour lequel les contributeurs sont à la fois les concepteurs et les bénéficiaires, comme c'est le cas de Wikipédia » (Néroutidis 2015). Cette distinction est également retenue dans des ouvrages de *Design Thinking* (Aldana, Dromer, Leméni 2019, par exemple). La définition de de Raphaëlle Bats n'est cependant pas invalidée, car elle se base sur une notion politique de démocratie qui va au-delà de ce que d'autres entendent par « participation ».

Il semblerait alors judicieux de retenir non pas une double, mais bien une triple distinction :

- Participation, au sens de « prendre part à » : on participe à un projet organisé par autrui, pour lequel une autorité désignée prend les décisions finales (sens *top-down*).
- Collaboration : on collabore avec les tenants du projet, et les décisions finales sont prises de manière démocratique (sens horizontal).
- Participation, au sens de « faire partie de » : les tenants du projet s'effacent pour laisser la place aux participants, qui deviennent décideurs à part entière (*bottom-up*).

Dans un service public, la participation peut prendre de nombreuses formes. Dans sa thèse portant sur la participation en bibliothèque, Raphaëlle Bats présente et compare deux modèles permettant d'évaluer le niveau de participation d'un projet : l'échelle d'Arnstein (fig. 4) et le Continuum proposé par les bibliothèques canadiennes du projet Working Together Libraries (WTL) (fig. 5) (Bats 2019).

Figure 4 : Échelle de la participation citoyenne selon Sherry Arnstein (1969)



(Bats 2019, p. 78)

Figure 5 : Continuum de la participation du public, du projet WTL (2008)



(Bats 2019, p. 80)

Si l'échelle d'Arnstein a pour avantage de présenter une forme d'évaluation des moyens de participation (le moins bon en bas et le meilleur en haut), le Continuum du projet WTL propose plutôt une forme de progression : toute forme de participation est bonne, mais on peut toujours essayer d'aller plus loin, de donner plus de place à l'utilisateur. Ce travail retiendra essentiellement ce second modèle. En effet, voir la participation comme un continuum horizontal facilite la gestion du changement non seulement pour les usagers mais aussi pour les équipes, car il permet une avancée progressive dans l'évolution des rôles de chacun, sans jugement de valeur.

Le Continuum du projet WTL donne en outre des exemples d'actions participatives possibles dans une bibliothèque, mais bien d'autres ont vu le jour. Les institutions patrimoniales, que l'on retrouve dans la littérature sous le nom de GLAM, pour *Galleries, Libraries, Archives, Museums* (Ridge 2014, par exemple), ne sont pas en reste, et des pratiques telles que la co-curation (Philippeaux 2015) et autres « Muséomix » (Chaumier, Françoise 2014) se répandent de plus en plus.

La diversité des pratiques découle de la variété des objectifs et des publics visés. En ce qui concerne la valorisation des collections patrimoniales numériques, une pratique participative courante et bien établie est celle du *crowdsourcing*.

3.1.1.2 Crowdsourcing

Le *crowdsourcing* a fait couler beaucoup d'encre, et l'on trouve d'excellents états de l'art récents qui traitent de l'un ou l'autre aspect de ce concept (Moirez, Moreux, Josse 2013 ; Néroutidis 2015 ; Dias da Silva, Heaton, Millerand 2017 ; Carron 2018 ; Boder, Dubosson 2020). Cette partie se contente de présenter les sources et éléments clés pour la compréhension du sujet et du présent travail.

Le terme *crowdsourcing* trouverait ses origines dans un article du magazine *Wired*, « The Rise of Crowdsourcing » (Howe 2006a). Dans un billet de blog publié le lendemain, l'auteur de ce mot-valise, composé de *crowd*, la foule, et d'*outsourcing*, externalisation, en donne cette définition :

« Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. » (Howe 2006b)

Si l'idée de base semble simple, ce concept a pourtant été interprété de bien des manières. Un article de 2012, visant à proposer une définition exhaustive du *crowdsourcing*, recense en effet une quarantaine de définitions différentes de ce terme, réparties entre trente-deux articles publiés entre 2006 et 2011 (Estellés-Arolla, González-Ladrón-de-Guevara 2012). À l'issue de leurs recherches, les auteurs proposent la définition suivante :

« Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken. » (Estellés-Arolla, González-Ladrón-de-Guevara 2012, p. 197)

Un ouvrage qui a fait date apporte une nuance à cette définition (Brabham 2013). L'auteur suggère en effet que les outils *open source* et les plateformes telles que Wikipédia ne peuvent pas être considérés comme du *crowdsourcing*, car l'organisme ne définit pas de lignes directrices à suivre ou d'objectifs à atteindre. Au contraire le management est laissé aux mains de la communauté (*bottom-up*). D'après lui, le *crowdsourcing* serait donc nécessairement participatif dans le sens *top-down*, et non collaboratif.

Cette distinction est reprise, mais affinée, dans un article de Trevor Owens. L'auteur y oppose *human computation* et *wisdom of the crowds*, la puissance de calcul humaine et la sagesse des foules. Son idée est de distinguer les projets de *crowdsourcing* qui s'appuient principalement sur la supériorité des capacités cognitives des individus par rapport aux ordinateurs, de ceux qui s'appuient avant tout sur l'intelligence collective (Owens, 2014). Selon sa définition, le *crowdsourcing* n'est pas forcément que participatif dans le sens *top-down*.

La distinction entre participatif et collaboratif est donc tout autant essentielle pour le *crowdsourcing* qu'elle ne l'est pour la définition de la participation en général. Il est alors intéressant de noter que, depuis 2014, on traduit en français le terme *crowdsourcing* par « production participative », mais que l'expression « production collaborative » est également acceptée (JORF 2014). En anglais cependant, on trouve une distinction triple entre *contributory projects*, *collaborative projects* et *co-created projects*. Dans le premier cas, les utilisateurs contribuent en produisant des données. Dans le second, ils aident également à analyser ces données et à améliorer l'organisation du projet, et, dans le dernier, ils sont directement impliqués dans sa conduite (Oomen, Aroyo 2011). Cette définition rejoint celle retenue pour la participation, et sera donc également retenue pour le *crowdsourcing*.

La complexité que représente la définition du *crowdsourcing* est entre autres liée à la grande diversité de projets de ce type. Certains sont à buts commerciaux, d'autres soutiennent des projets scientifiques – on parle alors de science citoyenne, et d'autres encore, et ce sont ceux qui nous intéressent, sont liés aux GLAM. Cette liste n'est bien sûr pas exhaustive.

3.1.1.3 Spécificités terminologiques et typologiques dans les GLAM

Dans le domaine patrimonial, le *crowdsourcing* prend des formes particulières. Un excellent état de l'art émanant de la Bibliothèque nationale de France, présente la participation dans ce contexte comme « la mise en œuvre de véritables compétences et connaissances des usagers, une interaction de haut niveau, de caractère scientifique », et le *crowdsourcing* comme des « projets collaboratifs de grande ampleur, mais l'accent sera davantage mis sur le nombre de participants, sur la notoriété du projet, sur la constitution de communautés de contributeurs, que sur la valeur scientifique de leurs contributions. » (Moirez, Moreux, Josse 2013, p. 6). Cette nuance repose à nouveau sur la place laissée aux usagers, mais aussi, et c'est essentiel, sur la légitimité accordée à leurs savoirs.

On note en outre l'importance de la notion de communauté. En effet, dans le contexte des GLAM, on parle volontiers de *community sourcing*, car les projets visent plus souvent des communautés d'usagers ou d'intérêt plutôt qu'une masse indistincte. Le terme *nichesourcing* est également utilisé pour des projets de niche, visant un public spécifique d'experts amateurs (Ridge 2013). La notion d'« amateur » prend ici une coloration particulière, et Trevor Owens souligne qu'il ne faut pas y voir de connotation négative, car Charles Darwin et Gregor Mendel étaient eux-mêmes des amateurs, au sens où ils n'étaient pas payés pour leur travail et leurs recherches (Owens 2014). Cette vision des « amateurs » sera retenue dans le présent travail.

La relation au public est donc au cœur de ce type d'initiatives, et est même souvent un objectif en soi. En parlant des musées, Graham Black insiste sur ce point :

« Thus, developing a participatory museum is not focused on increasing visitor numbers but, rather, on building lasting relationships between a museum and its audiences and/or communities and converting audiences into cultural participants, thereby increasing the museum's relevance to the communities it serves. » (Black 2018, p. 317)

Cette phrase, qui s'adapte à tous les GLAM, sera une ligne directrice de ce travail.

On retient trois grandes catégories de projets de *crowdsourcing* menés par des institutions patrimoniales :

- « Production de métadonnées dites « sociales », c'est-à-dire apport d'informations structurées qui enrichissent la description des documents ;
- Collecte de documents numérisés dont les originaux sont détenus par les utilisateurs ;
- Création de nouveaux savoirs, de textes scientifiques » (Moirez 2017, p. 131)

On entend par « métadonnées sociales » les données créées par les internautes et permettant d'améliorer la recherche et l'accès aux documents, ou d'offrir de nouveaux services aux usagers en répondant au mieux à leurs besoins (Moirez, Moreux, Josse 2013). Elles ne remplacent pas les métadonnées créées par des professionnels, bien entendu, mais elles viennent les compléter (Moirez 2013).

Ces métadonnées sociales peuvent prendre diverses formes : tags en vocabulaire libre, indexation en vocabulaire contrôlé, géolocalisation, correction d'OCR, transcription, traduction etc. Leur création n'a cependant de sens que si elles sont exploitées par la suite.

3.1.2 Recommandations et bonnes pratiques

On trouve, dans la littérature, des exemples de recommandations et de bonnes pratiques pour des projets de *crowdsourcing* dans le domaine patrimonial. Un article très influent, daté de 2010, est une mine d'informations en la matière. L'autrice y présente entre autres une checklist des éléments à prendre en compte pour réussir un projet.

Figure 6 : Checklist de Rose Holley

The Thing	The System	The People	The Content
👍 Clear goal	👍 Easy and Fun	👍 Acknowledge	👍 Interesting
👍 Big challenge	👍 Reliable and quick	👍 Reward	👍 New
👍 Progress	👍 Intuitive	👍 Team support	👍 Lots
👍 Results	👍 Options	👍 Trust	👍 History/Science

(Holley 2010, p. 9)

L'autrice de cet article présente également les bénéfices du *crowdsourcing* pour une institution culturelle, résumés dans la figure ci-dessous.

Figure 7 : Bénéfices du *crowdsourcing* pour des institutions culturelles, d'après l'article de Rose Holley (Holley 2010)

Objectifs	Communautés	Données	Relation aux usagers
<ul style="list-style-type: none"> • Les atteindre même si l'institution ne possède pas les ressources nécessaires • Les atteindre plus rapidement 	<ul style="list-style-type: none"> • Construire • Intégrer • Engager • Reconnaître et utiliser leurs connaissances, expertise et intérêt 	<ul style="list-style-type: none"> • Améliorer la qualité • Valoriser • Rendre les résultats de recherche plus précis • Améliorer la "trouvabilité" 	<ul style="list-style-type: none"> • Être plus à l'écoute et répondre mieux aux questions • Démontrer la valeur et la pertinence de l'institution • Renforcer la confiance et la loyauté envers l'institution • Encourager l'appropriation du patrimoine

La coopérative mondiale de bibliothèques OCLC propose également des recommandations, traduites en français dans la figure ci-dessous.

Figure 8 : Recommandations de l'OCLC pour des projets de *crowdsourcing*

- établir des objectifs clairs pour le projet : interactions avec les communautés d'utilisateurs ou enrichissement des collections ?
- motiver les usagers à contribuer et exploiter leur enthousiasme : identifier les raisons de contribuer (sujets intéressants, interfaces ludiques, contribution au bien commun, challenge, participation à une communauté)
- regarder d'autres projets pour glaner des idées avant de se lancer
- « Se lancer ! », ne pas se laisser intimider par la crainte des spam ou des malveillances : si les contributeurs sont bien encadrés, il y a peu de risques
- Mettre en place un règlement pour cadrer les interactions des usagers (aussi bien définition des comportements acceptables dans les commentaires par exemple, que définition des conditions légales de réutilisation des données produites par les usagers)
- Former et sensibiliser le personnel des institutions culturelles, à la fois en termes de technique et d'outils que de médiation avec les usagers
- Mettre en place des indicateurs pour mesurer le succès du projet (quantitatifs et qualitatifs)
- Analyser les atouts et les risques de déporter le projet sur un site tiers (par exemple Flickr)
- Utiliser de préférence des outils open-source
- Mettre en place des expérimentations avec les usagers avant et après le lancement du projet
- Ajouter régulièrement de nouveaux contenus, pour contribuer à maintenir l'intérêt de la communauté des contributeurs
- Rendre accessibles les données produites par les contributeurs, les indexer, les intégrer au catalogue
- Utiliser les projets de *crowdsourcing* pour construire de véritables communautés d'utilisateurs
- Utiliser des identifiants stables et pérennes pour désigner les objets numériques, afin de favoriser leur dissémination et leur visibilité
- Prévoir un plan de migration des contenus au cas où l'on voudrait changer de plate-forme
- Faire indexer les contenus par les moteurs de recherche (Google)
- Être réactifs, répondre rapidement aux demandes des usagers

(Moirez, Moreux, Josse 2013, p. 77)

Enfin, Donelle McKinley propose les principes suivants pour la création de projets de ce type, par ordre décroissant d'influence sur la participation et la qualité des contributions :

- « 1. *Provide clear, concise, and sufficient task instruction*
2. *Show how project output is freely accessible to the public*
3. *Keep the website current*
4. *Minimize the effort to contribute*
5. *Prioritize key information*
6. *Minimize user error*
7. *Enable users to review contributions*
8. *Clearly identify tasks*
9. *Present reasons to contribute*
10. *Provide task options*
11. *Simplify the task*
12. *Attractive design*
13. *Acknowledge participation*
14. *Encourage users to engage with the collection*
15. *Display project progress*
16. *Convey a sense of community*
17. *Convey the credibility of the project*
18. *Support community interaction*
19. *Publicly recognize contributions*
20. *Support content sharing*
21. *Convey a sense of fun »*

(McKinley 2015, p. 5)

Ces recommandations et bonnes pratiques sont fondamentales, et doivent être au cœur de la stratégie de tout projet de *crowdsourcing* par des institutions patrimoniales.

3.1.3 Publics et motivations

Les projets de *crowdsourcing* menés par des institutions culturelles attirent des usagers divers. L'article de Rose Holley, mentionné plus haut, propose quelques éléments concernant leurs profils (Holley 2010). Si certains profils semblent récurrents, chaque projet attire son public propre et développe sa communauté. Notons cependant qu'en règle générale, 90% des tâches sont effectuées par 10% des usagers (Carron 2018), d'où l'intérêt d'attirer dès le début une foule aussi nombreuse que possible.

La question des motivations demeure cependant cruciale, car il s'agit toujours de volontariat, et l'utilisateur doit pouvoir en tirer profit. Un article très récent (Alam, Sun, Campbell 2020) présente les motivations des usagers qui permettent d'initier la participation et de la maintenir. Les auteurs distinguent les motivations intrinsèques, donc liées à la pratique de l'activité pour elle-même, extrinsèques, liées aux incidences de cette pratique, et extrinsèques internalisées.

Figure 9 : Typologie des dynamiques de motivations pour le *crowdsourcing*

Motivation	Category	Initial Participation	Continued Participation
Intrinsic motivations	Personally (self)-oriented motivation	Own-use (self-interest), topic of interest	Learning, Addiction
	Affective motivation	Fun, Pastime, Passion for Topic, Simplicity, Task Autonomy	Fun, Pastime, Passion for Topic, Task Autonomy
	Community-based motivation	Altruism, Non-profit cause	Altruism, Kinship (Collectivism), Ownership
Internalised Extrinsic motivations	Self-esteem	Challenge	Challenge, Competition (internal self-concept)
	Pro-social motivation	Trust	Obligation/reciprocity
Extrinsic motivations	Reputation-based	Attribution	Recognition & Reputation-based reward
	Social Mechanisms	Indirect feedback	Advocacy

(Alam, Sun, Campbell 2020, p.17)

Comprendre ces motivations et les intégrer à la stratégie d'un projet de *crowdsourcing* est essentiel pour maintenir l'engagement des participants. Il demeure cependant utile de s'interroger sur les motivations spécifiques de la communauté propre au projet, pour pouvoir équilibrer, voire réorienter celui-ci.

3.1.4 Enjeux

Le mémoire d'Ariane Néroulidis propose une définition claire des enjeux juridiques, économiques et éthiques du *crowdsourcing* pour les institutions patrimoniales. On en retient en particulier l'importance de définir, dans les Conditions Générales d'Utilisation (CGU), si les données produites par les utilisateurs sont ou non protégées par le droit d'auteur ou autre législation, et de sélectionner des documents dont la diffusion et l'usage ne sont pas limités par de quelconques restrictions juridiques (Néroulidis 2015).

On note également que, d'un point de vue éthique, il faut prendre garde à ne pas tromper les usagers, en proposant du *crowdsourcing* « cosmétique » sans autre but que de « donner l'image d'une institution innovante » (Moirez 2013).

3.1.5 La situation helvétique

Dans les institutions patrimoniales suisses, le *crowdsourcing* demeure encore rare. Une étude de 2013 sur les institutions alémaniques a en effet démontré que des activités de volontariat en ligne n'ont été mises en place que dans 10% des institutions, et que 7% des institutions interrogées considèrent le *crowdsourcing* comme une opportunité à saisir (Estermann 2013).

Quelques initiatives ont porté leurs fruits, telles que celles de la bibliothèque de l'ETH (Wierkehr 2019) ou encore de la Fondation pour la sauvegarde du patrimoine audiovisuel de la RTS (Radio Télévision Suisse) et sa plateforme de partage et d'enrichissement de documents historiques publics et privés, notreHistoire.ch (Carron 2018). En ce qui concerne la correction d'OCR, un domaine auquel a ouvert la voie le projet australien TROVE (Moirez, Moreux, Josse 2013), on note l'initiative d'une plateforme de la Bibliothèque nationale, e-newspaperarchives.ch, dont les corrections d'OCR de journaux numérisés ont mobilisés plus d'une centaine de personnes, selon la responsable du projet, [Liliane Regamey](#). Il s'agit cependant de pratiques peu courantes dans le paysage helvétique.

3.2 Le cas des manuscrits

3.2.1 Le paléographe et la machine

Les documents manuscrits sont un cas particulier. Leur intérêt premier découle en effet de leur contenu écrit, qui doit être déchiffré pour pouvoir être traité. Contrairement aux documents imprimés, pour lesquels les logiciels OCR donnent des résultats généralement bons, tous les manuscrits ne peuvent encore être décryptés par une machine avec des résultats suffisants (Terras 2016). La plateforme [Transkribus](#), pionnière du domaine, repousse toujours plus loin les limites de la technologie, mais se révèle efficace essentiellement lors du traitement de manuscrits écrits de la même main. Une intervention humaine demeure donc nécessaire dans la majorité des cas pour produire des transcriptions fidèles à l'original.

Cet exercice, loin d'être aisé, est une discipline en soi que l'on nomme paléographie. Elle vise à produire des transcriptions qui pourront ensuite faire l'objet d'éditions, rendant l'accès au texte plus aisé et permettant d'en mieux comprendre la structure. Avec l'arrivée du numérique, le travail des paléographes et des éditeurs a évolué, en n'en devenant pas moins complexe.

L'usage de langages de balisage tels que le XML, recommandé par le *Text Encoding Initiative* (TEI), permet en effet de mieux rendre la structure d'un document (Vikhrova 2018). Dans sa thèse, Anne Vikhrova parle de « *decoding and encoding* » : on décode le manuscrit pour ensuite l'encoder en un langage compris par les machines (Vikhrova 2018, p. 20). Ce travail requiert néanmoins des compétences informatiques qui ne font pas partie du cursus de formation traditionnel des paléographes, mais dont les spécialistes en Humanités numériques sont les ambassadeurs (Terras 2016).

Une fois transcrits et édités, les manuscrits deviennent lisibles par des machines, permettant ainsi d'effectuer des recherches plein texte et facilitant leur « trouvabilité » (Severson, Sauvé 2019). En outre, cela rend le document intelligible pour des non-spécialistes. Il s'agit donc d'une étape cruciale pour qui veut faire connaître et rendre accessible ce type de documents.

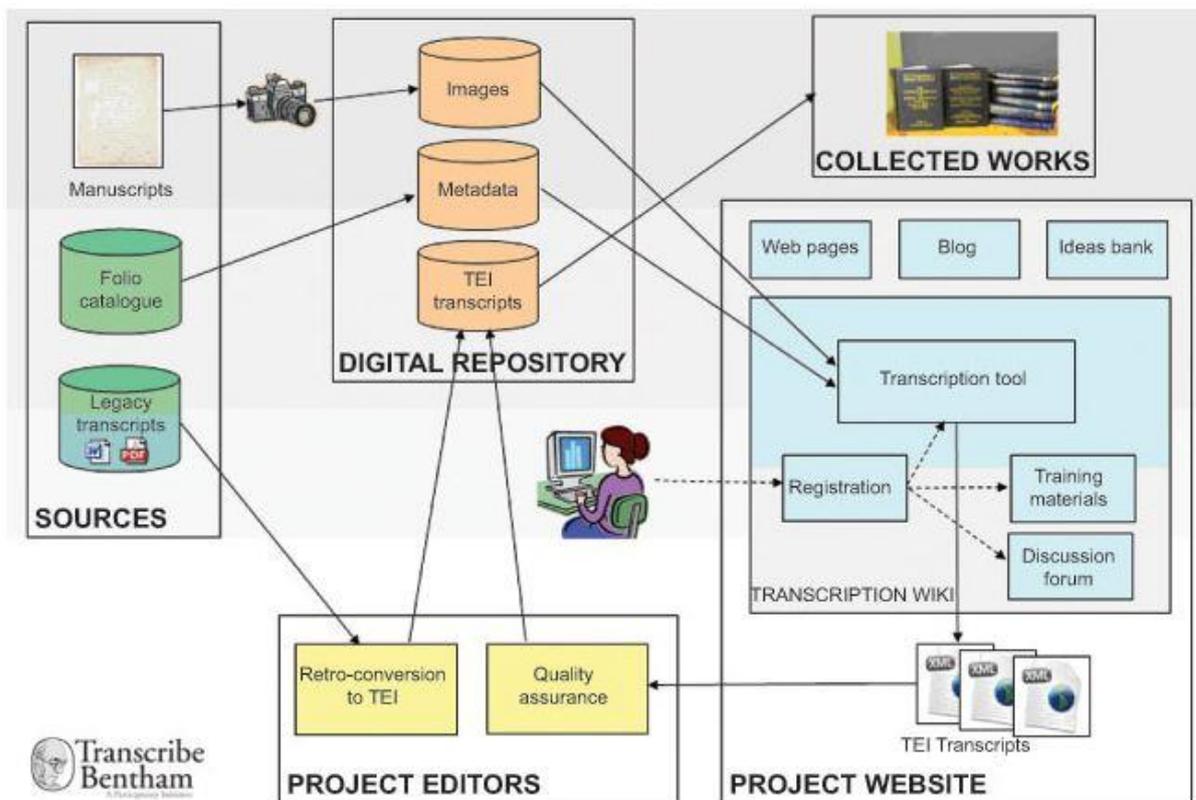
3.2.2 Une solution collective : la transcription collaborative

S'agissant d'un travail de longue haleine, il n'est pas surprenant que l'on ait cherché à recourir au *crowdsourcing* pour compléter le travail des chercheurs. La transcription collaborative est ainsi apparue comme la solution idéale.

Son fonctionnement de base est simple : un document numérisé, accompagné de ses métadonnées, est mis à disposition du public via un outil permettant de le visualiser et de transcrire son contenu, voire parfois d'y ajouter d'autres métadonnées (tags, par exemple). Ces transcriptions et métadonnées sont ensuite vérifiées et validées selon divers processus, puis intégrées à la base de données contenant ledit document, et dans lequel des usagers pourront faire des recherches. Ces recherches sont alors facilitées et améliorées grâce aux données produites par le public, les métadonnées sociales.

Ce *workflow*, généralement similaire pour la majorité des projets de transcription collaborative, a été détaillé et illustré par les tenants d'un projet de ce type, dont il sera question [plus loin](#).

Figure 10 : *Workflow* d'un projet de transcription collaborative



(Moyle, Tonra, Wallace 2011, p. 349)

On trouve bien entendu des différences de fonctionnement entre les projets. Par exemple, certains proposent de transcrire en texte brut, alors que d'autres donnent la possibilité de baliser ses transcriptions en XML. Ces balises doivent être parfois tapées manuellement ou ajoutées à l'aide de barres d'outils mises à disposition, selon la technique WYSIWYG, pour *What You See Is What You Get*, qui permet à l'utilisateur de voir directement le résultat de ses manipulations, et non leur version codée, plus complexe à comprendre. Le choix de l'une ou l'autre option découle directement de l'objectif visé. Si l'on souhaite simplement obtenir des transcriptions aisément lisibles par des humains et des machines, le texte brut peut suffire, mais si le but est de produire une édition numérique scientifique, qu'Anne Vikhrova nomme alors dans sa thèse l'édition érudite citoyenne, ou *Citizen Scholarly Editing* (Vikhrova 2018), des technologies plus complexes sont nécessaires.

Le processus de validation des transcriptions varie également. Dans un billet de blog daté de 2012, Ben Brumfield, une référence en matière de transcription collaborative, distingue deux grandes catégories de systèmes : *multi-track* et *single-track*. Le premier type implique qu'un même document est présenté à un certain nombre d'utilisateurs, et que, lorsque suffisamment de transcriptions coïncident, la version majoritaire est validée, avec ou sans l'intervention d'un expert, selon les cas. Le second type, au contraire consiste à ne proposer le document qu'en un seul endroit, et que les transcriptions puissent être faites par n'importe qui. Avec ce système, la transcription peut être laissée ouverte et modifiée à l'infini, avec un historique des versions, ou être fermée et validée par la communauté ou par un expert, sans possibilité de modification par la suite (Brumfield 2012a).

Ces distinctions dépendent le plus souvent du type de documents concernés. Les premières méthodes sont mieux adaptées à des documents structurés, *records*, en anglais, (registres, carnets de bord etc.) alors que les secondes correspondent mieux à du texte non-structuré (lettres, journaux intimes, manuscrits d'œuvres etc.) (Brumfield 2012a).

En plus de ces différences majeures, on note encore que certains projets donnent la possibilité d'ajouter des tags (géolocalisation, dates etc.), et intègrent des formes de réseaux sociaux, forums de discussion ou autres.

Toutes ces variations sont dues aux objectifs stratégiques du projet bien entendu, mais aussi aux outils utilisés. Ces outils sont extrêmement nombreux, comme le signale Ben Brumfield, lui-même développeur de l'outil *FromThePage*, avec sa femme Sara Brumfield (Brumfield 2013). Hormis un mémoire de Master de 2012, qui traite de six d'entre eux (Brokfeld 2012), aucune étude comparative ne semble avoir été publiée à ce jour. Un tel travail pourrait être d'une grande utilité à la recherche.

3.2.3 Projets existants

La transcription collaborative a fait ses preuves au fil de divers projets. Ne pouvant tous les mentionner, cette partie présente les projets fondateurs et « état de l'art » les plus pertinents pour le présent travail. D'autres listes sont cependant disponibles dans la littérature (Brumfield 2011 ; Alcalá Ponce de León 2015 ; Boder, Dubosson 2020, par exemple).

3.2.3.1 Projets fondateurs

Le projet de transcription collaborative le plus connu est sans doute [Transcribe Bentham](#). Lancé en 2010 par un groupe de chercheurs du University College London, ce projet a pour objectif principal d'établir une édition complète de Jeremy Bentham (1748-1832), sur la base des transcriptions produites par le public (Moyle, Tonra, Wallace 2011). S'inspirant des bonnes pratiques de Rose Holley (Holley 2010), leur but est également de créer une communauté et d'attirer l'attention du public vers la pratique de la transcription de manuscrits ainsi que vers les œuvres méconnues de Bentham. Ce projet utilise un outil *open source* développé en interne, *Transcription Desk*, basé sur MediaWiki et permettant le balisage des transcriptions en XML par les contributeurs (Moyle, Tonra, Wallace 2011). Le projet n'est pas encore fini mais, rien qu'entre 2012 et 2014, plus d'un million de mots ont été transcrits, dont à peine 1% ont nécessité des modifications par l'équipe de chercheurs (Causer et al. 2018). **Transcribe Bentham** est un succès indiscuté qui a permis de faire connaître la transcription collaborative au grand public (Brumfield 2011).

Un autre projet que l'on peut considérer comme « fondateur » est [Old Weather](#), lancé en 2010 par le service météorologique et le Musée Maritime National du Royaume-Uni, le site Naval-History.net et la plateforme de science citoyenne Zooniverse. Il utilise l'outil de *crowdsourcing* de cette dernière, devenu depuis *open source* sous le nom de *Scribe*. Présenté à l'origine comme un projet de science météorologique, il a été considéré comme un projet de recherche historique par son public qui l'a traité ainsi, et c'est cet aspect qui aurait finalement attiré et fait rester les volontaires (Blaser 2014).

La New York Public Library est également à l'origine de projets qui ont fait date. Le plus célèbre d'entre eux, [What's on the menu ?](#), lancé en 2011, a permis la transcription d'environ 9'000 images de menus de restaurants numérisés (Lascarides, Vershbow 2014). Le sujet,

quoiqu'étonnant, a pourtant attiré les foules, que ce type de documents intriguaient et passionnaient, et le projet a été couronné de succès.

The Civil War Diaries and Letters Transcription Project, lancé en 2011 par la bibliothèque de l'Université d'Iowa et à l'origine de la plateforme [DIY History](#), est également un projet à succès. Dans ce cadre a été développé l'outil *open source* *Scripto* (Leon 2014), repris depuis par des projets d'humanités numériques tels que [Transcrire](#), par exemple.

Enfin, **DigitalKoot**⁵, un projet finlandais lancé en 2011, a pour spécificité d'être « gamifié ». Au lieu de simplement transcrire des documents, l'internaute joue à des jeux qui nécessitent de transcrire correctement des fragments de texte pour gagner (Chronos, Sundell 2011).

3.2.3.2 Projets récents et « état de l'art »

L'un des projets récents qui a fait beaucoup parler de lui est le [Transcribathon](#) d'Europeana, qui donne accès non seulement à des manuscrits mais aussi à des objets, photographies et autres, que les volontaires peuvent tagger. Ces documents sont issus de multiples institutions, l'un des objectifs d'Europeana étant de donner accès aux collections patrimoniales européennes en les agrégeant (Europeana 2020). Ce projet se construit comme un concours, et l'utilisateur reçoit des points, nommés *miles*, à chacune de ses contributions, ce qui lui permet d'accéder à plus de fonctionnalités (relecture et validation, par exemple).

La plateforme Zooniverse continue de proposer des projets de transcriptions. Parmi les plus récents, on trouve [Anti-Slavery Manuscripts](#), lancé par la Boston Public Library, et [Every Name Counts](#), un projet des Arolsen Archives visant à identifier les détenus de camps de concentration nazis.

Plusieurs grandes institutions ont récemment monté leurs propres projets de transcription collaborative. Non des moindres, la Library of Congress (LoC) a lancé, en 2018, [By the People](#), un projet basé sur *Concordia* un outil *open source* développé en interne avec la méthode Agile, (Ferriter et al. 2019). La British Library a également développé son propre outil *open source* basé sur *Pybossa*, [LibCrowds](#), pour sa plateforme du même nom.

Enfin, la thèse d'Anne Vikhrova, dans laquelle elle proposait un prototype d'outil d'édition érudite citoyenne (Vikhrova 2018), a donné lieu à la création de la plateforme [TACT](#) et de son outil de transcription *open source* du même nom.

3.2.3.3 Projets en Sciences naturelles et en Histoire de l'art

Moins courants, les projets de transcription collaborative en Sciences naturelles ou en Histoire de l'Art, dont trois exemples sont présentés ici, valent néanmoins le détour. [Les Herbonautes](#) est un projet très réussi de transcription d'étiquettes d'herbiers lancé en 2012, dont le but est de permettre une meilleure documentation des spécimens en enrichissant les bases de données existantes (Chupin 2016). [North American Bird Phenology Program](#), lancé en 2009, est un projet dont le but, atteint en 2016 (Brumfield 2020), était d'obtenir la transcription d'environ 6 millions de cartes d'observations d'oiseaux (Terras 2016). Enfin, le projet pilote de [Co-Lab](#), l'espace de transcription des Bibliothèque et Archives du Canada, portait sur la transcription d'étiquettes d'anciens objets en argent québécois (Severson, Sauvé 2018).

⁵ Ce site n'est plus disponible.

3.2.3.4 Projets suisses

Comme indiqué [ci-dessus](#), les projets de *crowdsourcing* qui ont vu le jour en Suisse sont rares. Dans le domaine de la transcription collaborative, on note cependant deux projets en particulier. L'EPFL a en effet lancé en 2019 sa bibliothèque numérique [PLUME](#), qui comporte un outil d'annotations des documents permettant aussi la transcription (Granval, D'Ans 2019). Cette plateforme a été élaborée avec l'aide de la société de numérisation et de valorisation patrimoniale Arkhênum, qui propose une checklist pour réussir un projet de ce type (Arkhênum [sans date]). En Suisse alémanique, la plateforme [e-manuscripta.ch](#), fondée par la bibliothèque de l'ETH, la Zentralbibliothek Zürich et l'Universitätsbibliothek Basel, propose depuis fin 2018 un outil de transcription des collections (Baumann 2018). Ces projets n'ont cependant pas encore reçu beaucoup d'attention de la part du grand public.

4. Résultats et analyse

Ce chapitre présente les résultats des différentes étapes de recherche.

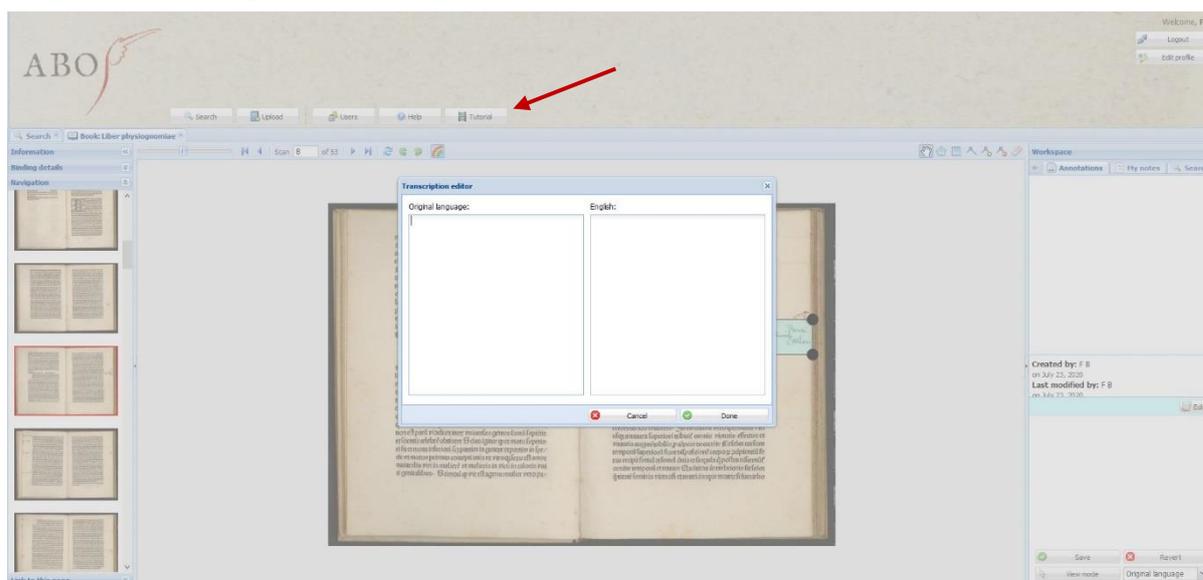
4.1 Observations participantes

La [grille d'expérience](#), réalisée au préalable, a permis de tester dix plateformes de transcription existantes, aussi différentes que possibles et basées sur des outils variés. Certaines d'entre elles ont été mentionnée [ci-dessus](#), et les éléments présentés ne sont en principe pas répétés. Pour chacune, on a noté les bonnes surprises et les éléments irritants ou décourageants, selon l'expérience vécue. Une nuance est parfois apportée dans le cas où ces éléments semblent découler du type de public ciblé par le projet.

4.1.1 Annotated Books Online

[Cette plateforme](#), gérée par des universités anglo-saxonnes et néerlandaises, réunit des imprimés anciens comportant des annotations manuscrites, dont elle propose la transcription. Elle s'adresse avant tout à des chercheurs et étudiants universitaires. L'inscription est obligatoire, et la relecture est faite par les autres usagers selon un modèle *single-track*. L'outil de transcription est *open source* et a été créé spécifiquement pour la plateforme en question.

Figure 11 : Capture d'écran de Annotated Books Online



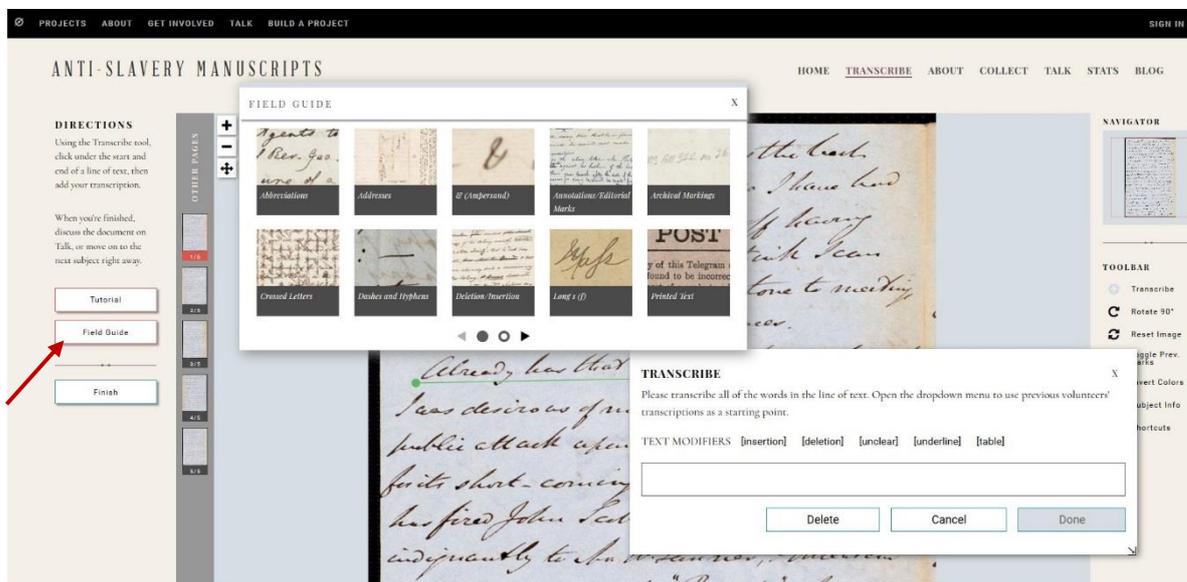
- | | |
|---|--|
| ✓ Tutoriel visible et clair | ✗ Apparence ni moderne ni attrayante |
| ✓ Possibilité de sélectionner une zone de texte pour cibler les annotations | ✗ Transcription « pop-up » qui cache parfois le texte |
| ✓ Possibilité de traduire aussi | ✗ Pas d'indication du niveau d'avancement des transcriptions |

Ces points découlent essentiellement du fait que la plateforme a des objectifs et un public précis, et n'a pas de vocation « grand public ». Le système de pop-up pour la transcription peut néanmoins être contrariant à la longue.

4.1.2 Anti-Slavery Manuscripts

[Ce projet](#), hébergé sur la plateforme Zooniverse, a pour but d'obtenir la transcription de correspondances manuscrites d'activistes abolitionnistes du XIX^{ème} siècle. L'inscription n'est pas obligatoire, et la validation suit un modèle *multi-track*.

Figure 12 : Capture d'écran de Anti-Slavery Manuscripts

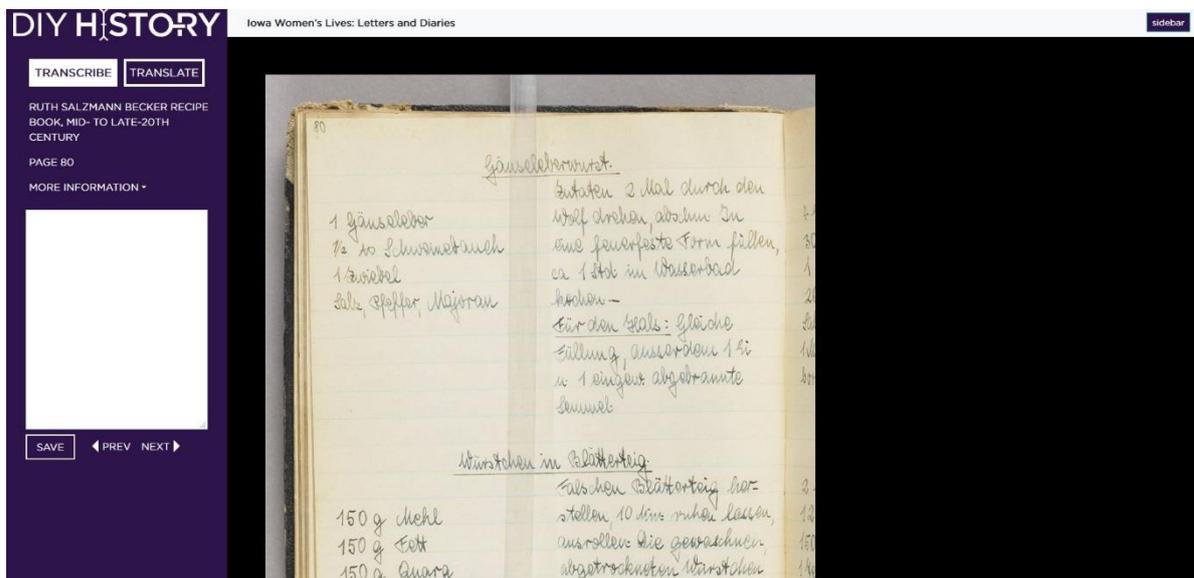


- ✓ Tutoriel pop-up clair
- ✓ « Field Guide » très bien fait
- ✓ Raccourcis pour les balises XML (WYSIWYG)
- ✗ Nécessité de souligner chaque ligne de texte
- ✗ Pas de choix du document
- ✗ Transcription « pop-up » qui cache parfois le texte

4.1.3 DIY History

[Cette plateforme](#) est gérée par l'Université d'Iowa et fonctionne avec l'outil *Scripto*. L'inscription est obligatoire, et la validation suit un modèle *single-track*.

Figure 13 : Capture d'écran de DIY History

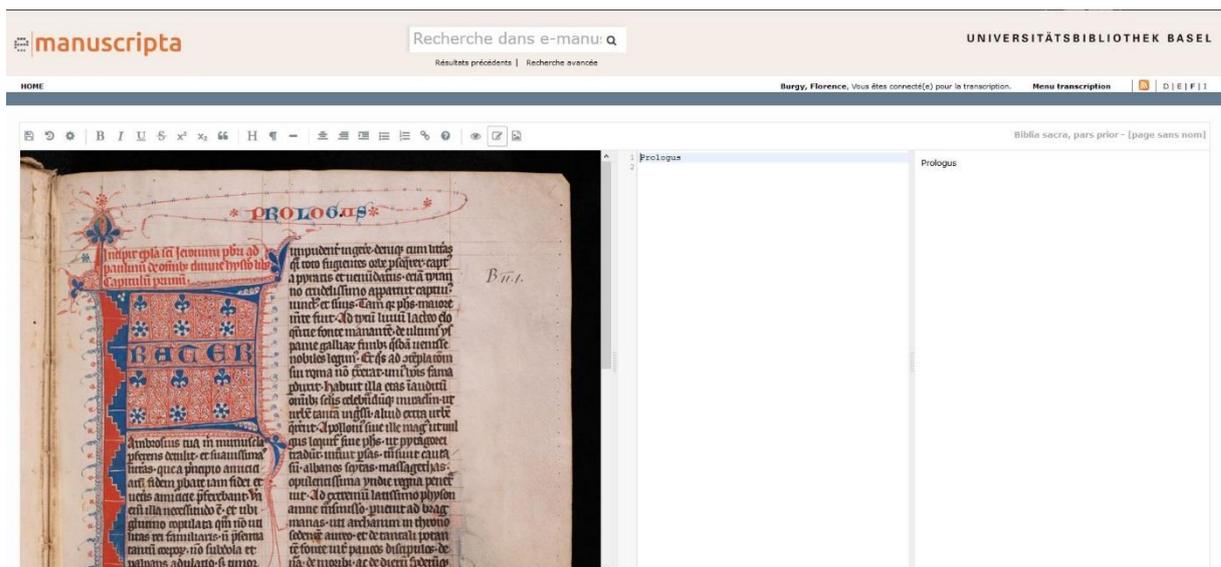


- ✓ Multiples projets clairement distincts
- ✓ Choix de documents, avec indication de l'état d'avancement
- ✓ Possibilité de traduire aussi
- ✗ Pas de tutoriel, peu d'informations en général
- ✗ Règles de transcription sur une autre page : risque de perdre sa transcription si l'on s'y rend
- ✗ Transcription considérée comme terminée une fois sauvegardée, même si un ou deux caractères seulement ont été inscrits

4.1.4 E-manuscripta

[Cette plateforme](#) présente les collections de manuscrits de diverses institutions alémaniques, qu'il est possible de transcrire grâce à un outil d'annotation développé en interne. L'inscription est obligatoire, et la relecture suit un modèle *single-track*, avec une validation effectuée par les institutions en charge du projet.

Figure 14 : Capture d'écran de e-manuscripta



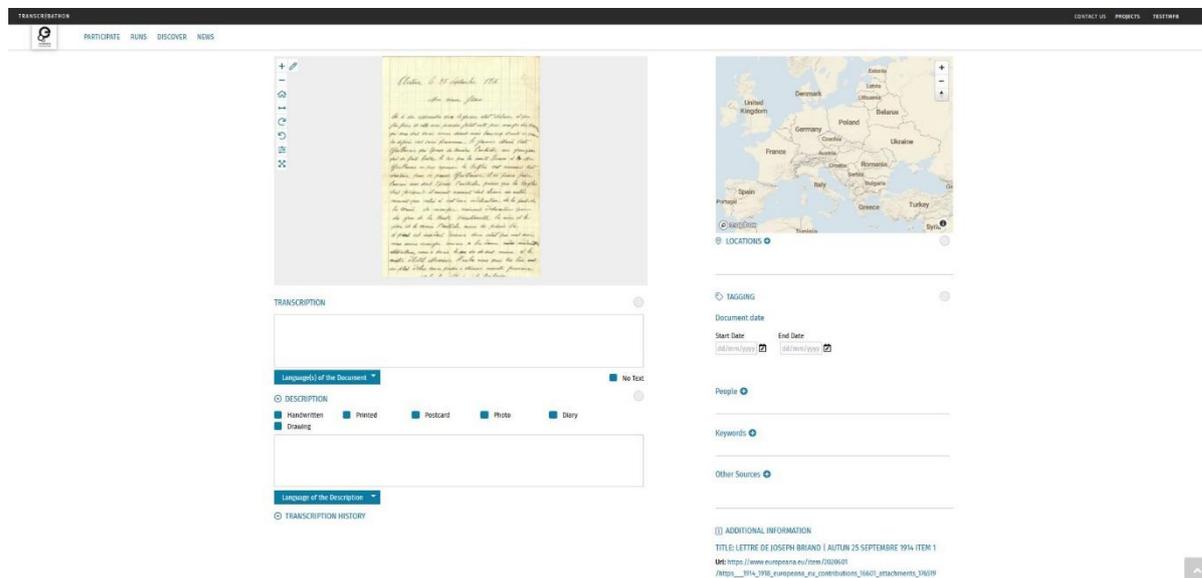
- ✓ Mode d'emploi et règles de transcription clairs
- ✓ Possibilité de choisir les documents
- ✓ Outil aisé d'utilisation, avec formatage en Markdown
- ✗ Trop de choix : plus un catalogue qu'un projet
- ✗ On ne sait pas ce qui a déjà été transcrit
- ✗ Règles de transcription sur une autre page : risque de perdre sa transcription si l'on s'y rend

Comme pour Annotated Books Online, ces points pourraient *a priori* découler du type d'utilisateurs ciblé, et du fait qu'il ne s'agirait pas d'un projet « grand public ». Cela n'est cependant pas très clair ici, et il est difficile de comprendre à qui s'adresse ce projet.

4.1.5 Europeana – Transcribathon

[La plateforme](#) de *crowdsourcing* d'Europeana, développée en interne, nécessite une inscription. La relecture et validation sont effectuées par les autres usagers, selon un modèle *single-track* particulier, basé sur un système de points et de niveaux.

Figure 15 : Capture d'écran d'Europeana - Transcribathon

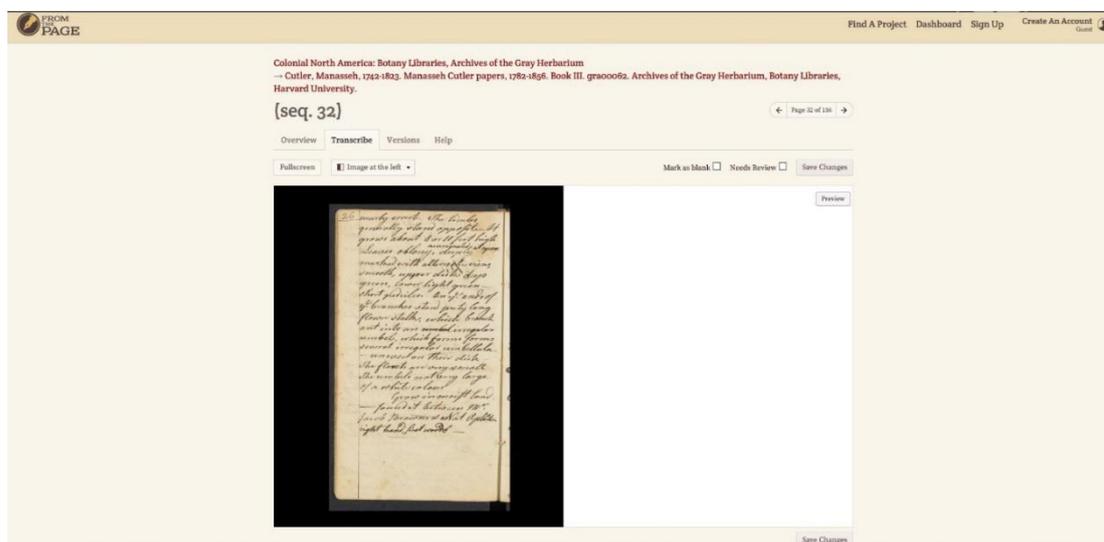


- ✓ Choix de documents avec indication de l'état d'avancement
- ✓ Outil de transcription facile d'utilisation
- ✓ Plusieurs possibilités d'enrichissements
- ✗ Temps de chargement long
- ✗ Système d'inscription/login pas clair et peu visible
- ✗ Tutoriel et règles de transcription peu aisées à trouver

4.1.6 Harvard Library – Archives of the Gray Herbarium

[Ce projet](#) est hébergé par la plateforme FromThePage, dont elle utilise l'outil de transcription. L'inscription n'est pas obligatoire, mais une limite est mise au nombre de transcriptions en anonyme. La relecture suit un modèle *single-track*.

Figure 16 : Capture d'écran du projet Archives of the Gray Herbarium

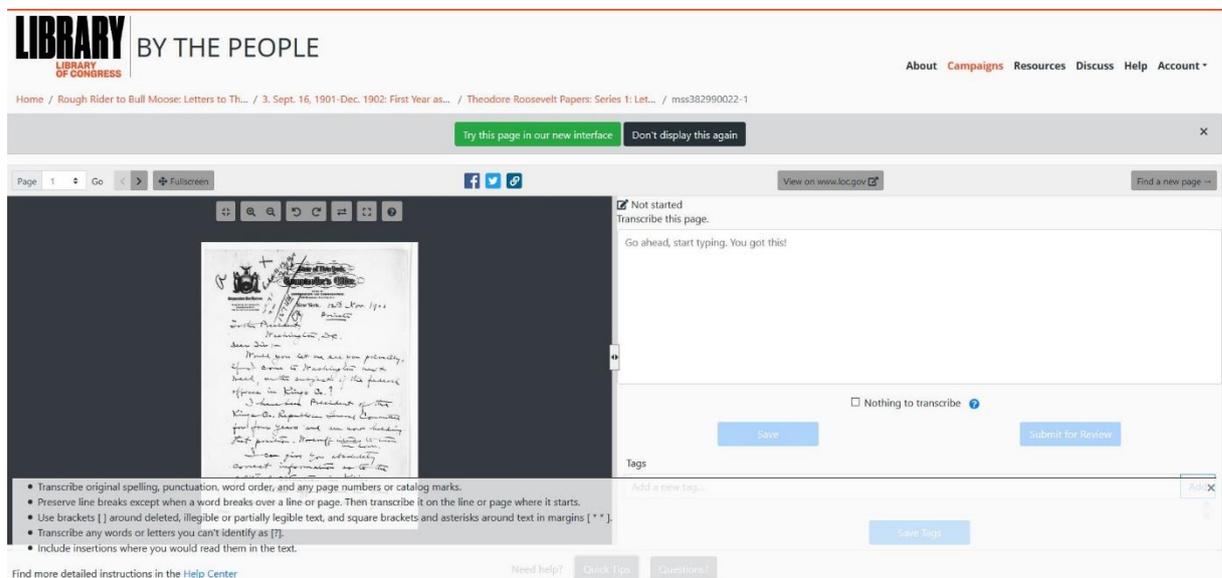


- ✓ Choix de documents avec indication de l'état d'avancement
- ✓ Règles de transcription sous l'image
- ✓ Possibilités de placer la zone de transcription à côté de, sur ou sous l'image
- ✗ Apparence peu moderne
- ✗ Pas de didacticiel
- ✗ Un peu difficile de s'y retrouver dans la liste des documents

4.1.7 Library of Congress : By the people

[La plateforme](#) de la Library of Congress utilise un outil développé en interne, *Concordia*. L'inscription n'est pas obligatoire, mais il faut un compte pour pouvoir relire et tagger. Le mode de validation de cette plateforme semble être récemment passée d'un système *multi-track* à un système *single-track* (entre le test et la rédaction de ce travail).

Figure 17 : Capture d'écran de Library of Congress : By the People

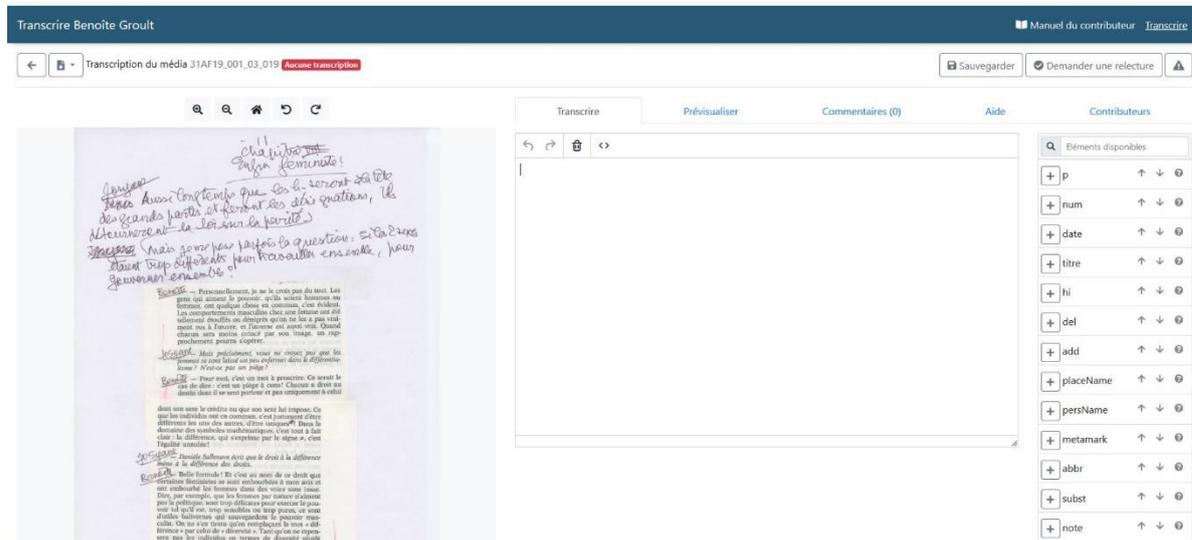


- ✓ Apparence moderne, attrayante
- ✓ Instructions claires et immédiatement visibles
- ✓ « Quick Tips » rappelant les éléments importants du processus de transcription
- ✗ Pas de choix de document (ceci a changé depuis le test)
- ✗ Règles de transcription sur une autre page : risque de perdre sa transcription si l'on s'y rend
- ✗ Pas de possibilité de formatage du texte

4.1.8 TACT

La plateforme française [TACT](#) utilise l'outil *open source* du même nom, et nécessite une inscription. Certains projets ne sont accessibles que sur demande. La relecture suit un modèle *single-track*, et la validation revient aux tenants du projet.

Figure 18 : Capture d'écran de TACT



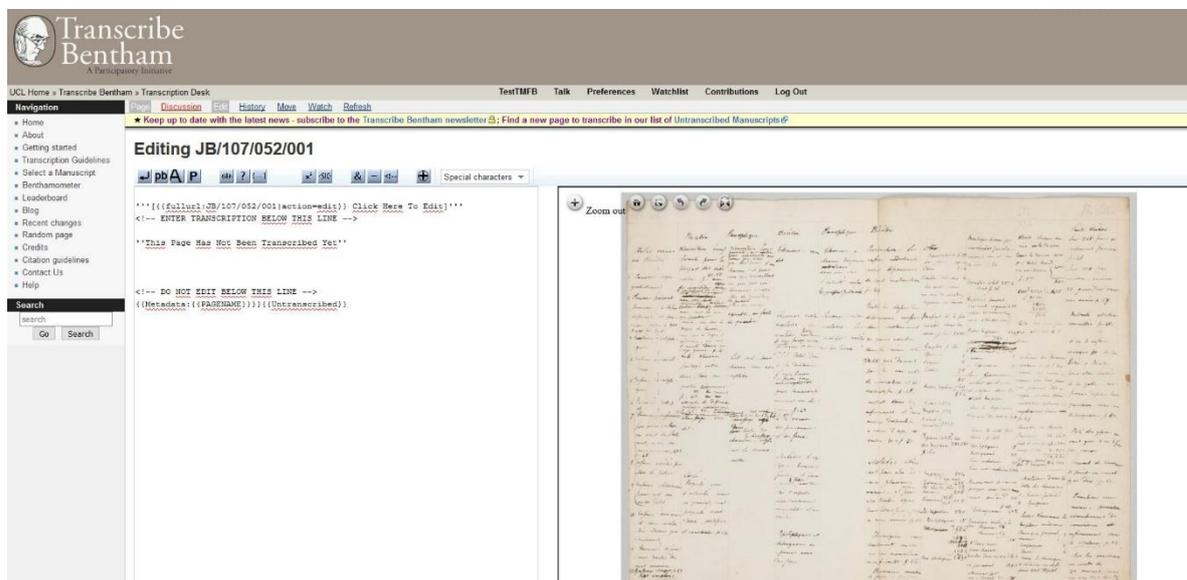
- ✓ Manuel clair
- ✓ Choix de documents avec indication de l'état d'avancement
- ✓ Possibilité de voir les documents transcrits par autrui, pour mieux comprendre ce qu'on attend de nous
- ✗ Choix de document folio par folio
- ✗ Les manuels ne sont accessibles qu'en PDF
- ✗ Les balises ne sont pas évidentes à utiliser et compliquent la tâche

Comme Annotated Books Online et e-manuscripta, le public ciblé par la plateforme n'est pas très clair. Selon les projets, il semble que cela va du « grand public » aux étudiants inscrits à un séminaire universitaire en particulier. Cependant, ici, les projets et les fonds concernés sont distincts, et les modalités d'accès à chaque projet sont précisées.

4.1.9 Transcribe Bentham

La célèbre plateforme [Transcribe Bentham](#), basée sur l'outil *Transcription Desk*, nécessite une inscription pour contribuer. La relecture suit un modèle *single-track*, et la validation est effectuée par les responsables du projet.

Figure 19 : Capture d'écran de Transcribe Bentham

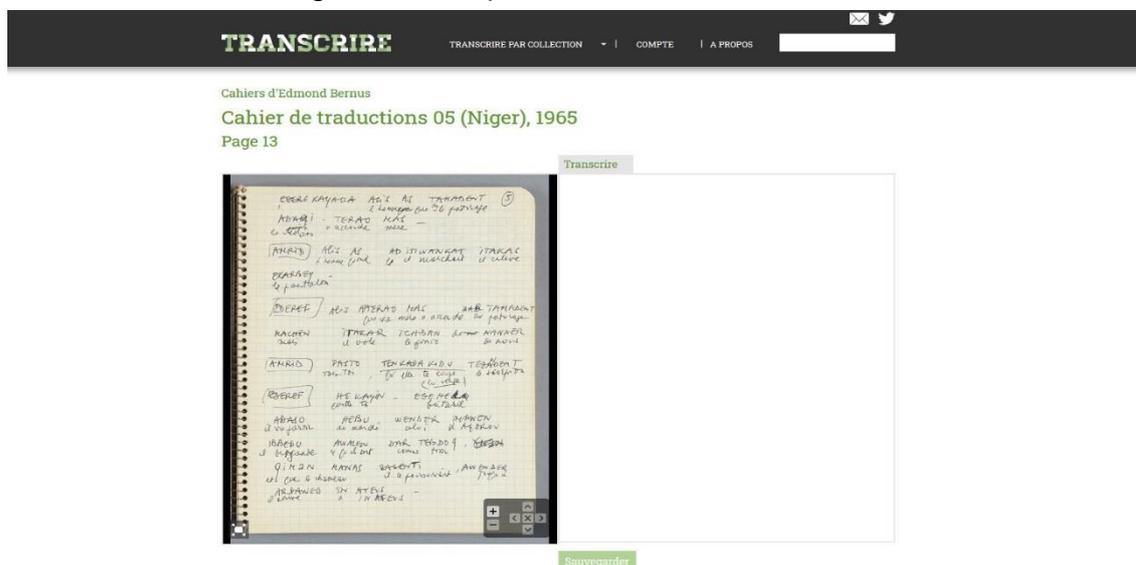


- ✓ Projet bien documenté
- ✓ Instructions très détaillées
- ✓ Choix de documents avec différents niveaux de difficulté
- ✗ Cotes comme seuls intitulés
- ✗ Règles de transcription sur une page à part
- ✗ Les balises ne sont pas évidentes à utiliser et compliquent la tâche

4.1.10 Transcrire

Le projet français d'Humanités numériques [Transcrire](#) utilise *Scripto*, comme la plateforme DIY History, mais se présente assez différemment. L'inscription est obligatoire, et la relecture est *single track*, mais la validation est effectuée par les responsables de projet.

Figure 20 : Capture d'écran de Transcrire



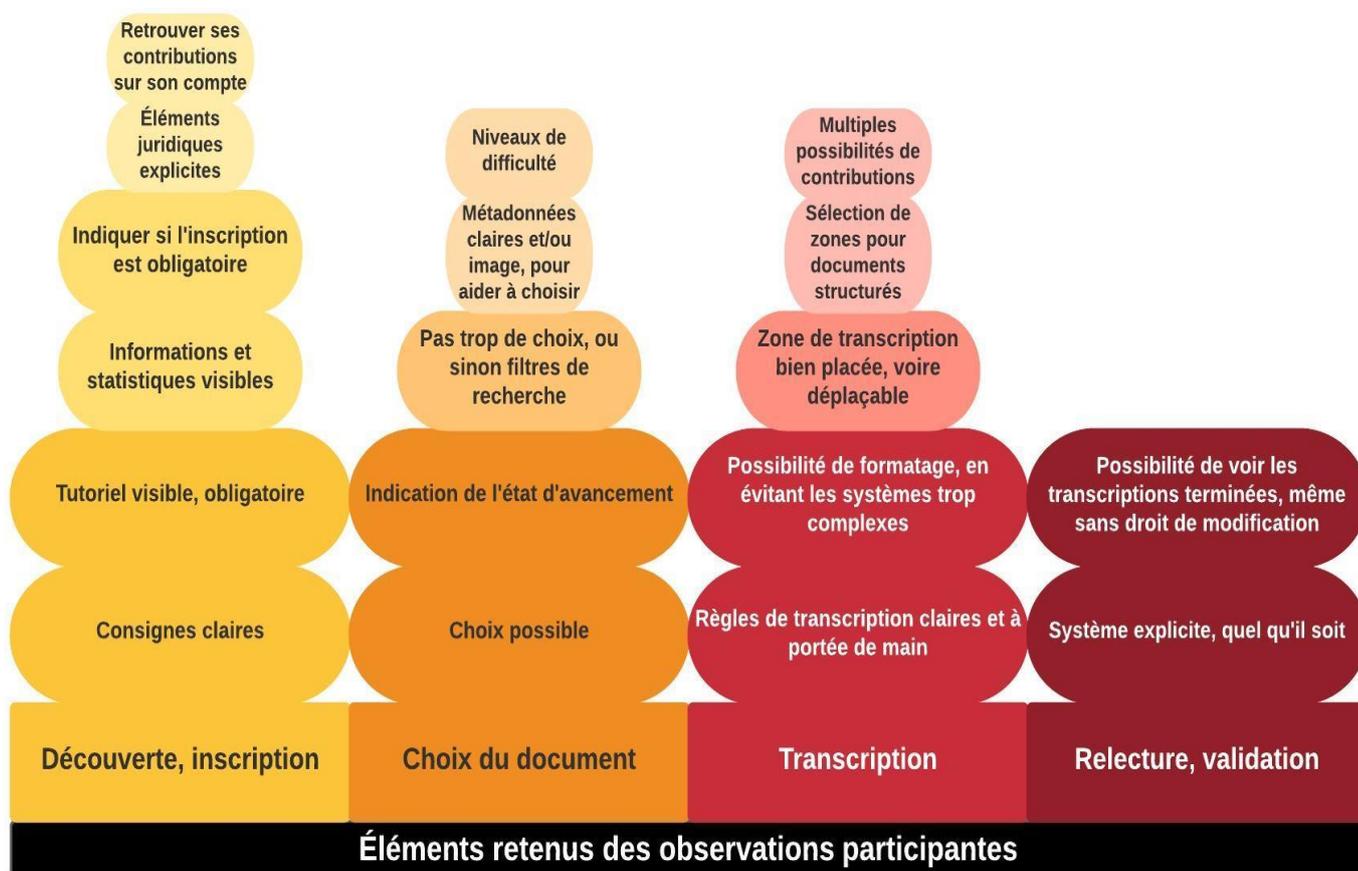
- ✓ Apparence moderne, attrayante
- ✓ Choix de documents avec indication de l'état d'avancement
- ✓ Outil de transcription plus ergonomique que sur DIY History
- ✗ Pas vraiment de didacticiel
- ✗ Règles de transcription sur une autre page : risque de perdre sa transcription si l'on s'y rend
- ✗ Pas de possibilité de formatage du texte

4.1.11 Éléments à retenir des observations participantes

Les différents tests ont permis de noter certains éléments encourageants ou décourageants, qui peuvent avoir une influence sur l'expérience de la transcription collaborative. Dans la figure ci-dessous, ces éléments sont classés en quatre catégories, selon le moment où ils interviennent dans le processus de transcription, ainsi que par degré d'importance (les plus grandes bulles étant les plus importantes, bien entendu).

Ces éléments découlent néanmoins uniquement d'une expérience personnelle, et il est essentiel de nuancer cette analyse avec les retours d'expérience des [responsables de Mirabilia](#) et des [usagers potentiels](#), ce qui sera fait par la suite.

Figure 21 : Éléments retenus des observations participantes



4.2 Entretiens avec des experts

Contacter des experts permet de récolter des conseils et retours d'expérience, mais aussi de nouer des liens pour des partenariats futurs. Les personnes retenues sont soit expertes dans le domaine des manuscrits, soit des personnes responsables de projet de *crowdsourcing* en Suisse. Le département d'Humanités numériques de l'UNIGE a également été contacté, mais ce type de projet visant des amateurs sort de leur champ d'action. Le Bodmer Lab, pionnier genevois en matière d'Humanités numériques, a également été contacté, sans succès, mais il demeure un partenaire potentiel intéressant. Les guides d'entretien qui ont servi à préparer ces rencontres sont disponibles en [annexe](#), et toutes les personnes concernées ont donné leur accord écrit pour la diffusion de leurs propos, qu'elles ont elles-mêmes validés.

4.2.1 David Glaser – responsable de [notrehistoire.ch](#) – 16.04.2020

La plateforme [notrehistoire.ch](#) ne propose pas de transcription collaborative de manuscrits, mais il s'agit néanmoins d'une référence suisse en matière de valorisation du patrimoine numérisé et de *crowdsourcing*.

Cet entretien avec David Glaser a d'abord été l'occasion de discuter des résultats d'une enquête menée auprès des usagers de ladite plateforme (Glaser 2020). Le rapport du sondage permet de relever que les usagers sont attirés par la qualité des documents proposés et la variété des fonds (publics et privés), avec la possibilité de se divertir et l'ancrage local (romand), en lien avec une institution importante (la RTS), comme points forts. Le public n'est cependant pas prêt à payer pour un tel service, et se soucie du fait que ses données ne soient pas revendues. Les usagers attendent de la nouveauté et de la simplicité, et plébiscitent l'ouverture à tous. Les contenus semblent donc importer plus que la modération et l'animation, qui demeurent néanmoins essentielles.

M. Glaser indique en effet que la modération témoigne de la qualité et de la légitimité de la plateforme, mais qu'elle doit rester discrète, la présence de l'équipe éditoriale de manière permanente et active – sur la page d'accueil – n'étant pas ce que cherchent en priorité les personnes interrogées lors du sondage. Il faut trouver un bon équilibre.

Les rencontres physiques entre les membres ne semblent pas importer non plus, mais un cinquième des répondants indique tout de même être intéressé par des événements apportant des compléments d'information par rapport à la plateforme (conférences, débats etc.). En revanche les espaces de communication en ligne sont appréciés des usagers, avec la modération qui s'impose.

Concernant les facteurs d'échec et de réussite, M. Glaser insiste sur l'importance de la communication et de la définition des publics cibles, comme le monde de l'éducation et de l'instruction publique, ou les sociétés savantes. Un tel projet doit se faire connaître auprès des bonnes personnes. Il recommande en outre de commencer en « frappant fort », avec un projet charismatique, qui marque les esprits et donne envie de participer.

La plateforme Mirabilia intrigue M. Glaser, qui fait part de sa disponibilité pour toute question concernant les bonnes pratiques et risques d'échecs qui découlent de son expérience, et note qu'un partenariat, ou du moins un échange de bons procédés, avec [notrehistoire.ch](#) est tout à fait envisageable.

4.2.2 Barbara Roth – conservatrice responsable des collections spéciales de la BGE récemment retraitée – 04.05.2020

Barbara Roth, ancienne responsable des collections spéciales de la Bibliothèque de Genève, demeure une référence en matière de conservation et de traitement des manuscrits. Ses collègues et successeurs représentent cette institution au sein du Comité éditorial de Mirabilia, et Mme Roth nous invite à se référer à eux avant tout, mais elle accepte volontiers de nous faire part de ses conseils et de son expertise.

L'entretien porte d'abord sur quelques points précis. En ce qui concerne les publics à cibler, le « grand public » semble compliqué à atteindre et, d'après Mme Roth, mieux vaut cibler des communautés et groupes spécifiques. Elle mentionne entre autres des groupes de recherche sur François Tronchin ou Ferdinand de Saussure, mais aussi le milieu éducatif, ainsi que les migrants découvrant la culture genevoise. En matière de fonds à valoriser, elle conseille de prioriser les fonds les plus consultés, pour lesquels une demande existe déjà, mais recommande d'attirer aussi l'attention sur des fonds riches mais moins connus. Enfin, concernant les règles de transcription, Mme Roth conseille de se référer aux normes de l'École des Chartes. Le choix des règles dépend cependant du public et des objectifs visés.

Concernant le projet dans son ensemble, Mme Roth note qu'un dialogue entre les institutions et les chercheurs est essentiel. Les chercheurs en savent souvent plus que les spécialistes de l'information sur certains documents, et les projets de transcription peuvent aussi partir d'eux. L'expertise des spécialistes de l'information demeure néanmoins nécessaire, et une collaboration des deux parties pour l'élaboration du projet et pour le soutien aux futurs usagers semble être une bonne voie à suivre.

Mme Roth souligne en outre l'importance de l'interopérabilité des données et des outils dans leur ensemble. Trop souvent, les institutions suisses tendent à mettre sur pied des projets « dans leur coin », alors que des collaborations, notamment avec les nouveaux spécialistes des Humanités numériques, seraient profitables à tous. Elle met également en garde contre les projets « prétextes », qui servent à donner une certaine image de l'institution, sans autre réel objectif stratégique ou bibliothéconomique, et qui, de ce fait, peuvent tromper les usagers.

4.2.3 Liliane Regamey – responsable de e-newspaperarchives.ch – 25.05.2020

Liliane Regamey est responsable de la section Utilisation de la Bibliothèque nationale suisse (BN). Dans ce cadre elle est entre autres en charge de la plateforme de valorisation de journaux suisses numérisés, e-newspaperarchives.ch, pour laquelle un projet de correction d'OCR a été élaboré. Comme pour M. Glaser, cette rencontre a pour but de mieux connaître le paysage suisse en matière de *crowdsourcing*, et de récolter conseils et retours d'expérience.

Concernant ce projet, Mme Regamey note qu'au début, peu d'utilisateurs de la plateforme utilisaient l'outil de correction. Un projet spécifique a alors été élaboré, en lien avec les 60 ans du premier refus du droit de vote des femmes en Suisse, en novembre 1959. Les journaux de la plateforme, retraçant les débats de l'époque, étaient à disposition pour être corrigés et taggés en langage naturel. Le thème proposé était une suggestion et les participants avaient la possibilité de traiter tous les articles des journaux ouverts à la correction.

La campagne de communication a été commencée longtemps à l'avance afin de préparer le terrain. Comme ce thème visait un public généraliste, intéressé par la presse, l'Histoire suisse et l'Histoire des femmes, le marketing est passé par divers canaux : réseaux sociaux, site web

de la BN, liste de diffusion swiss-lib, communiqué de presse, ainsi qu'une intervention lors de l'émission de radio Médialogues. Cette campagne a permis de recruter 129 participants, mais c'est l'intervention à la radio qui semble avoir eu le plus gros impact, car 76 personnes se sont inscrites dans les jours qui ont suivi l'émission, donc plus de la moitié. Mme Regamey souligne alors que les médias traditionnels ne doivent pas être négligés, bien au contraire

Afin de rendre le projet plus motivant, il a été construit comme un défi limité dans le temps, lors duquel le nombre de contributions des usagers était comptabilisé. À la fin de l'opération, les trois meilleurs participants se voyaient invités à la BN pour une visite privée, ce qui était une motivation pour certains.

Concernant la modération, Mme Regamey explique qu'elle demandait un temps peu considérable, comme aucun système de communication entre les usagers n'avait été installé. La vérification des corrections se faisait par échantillonnage, pour s'assurer qu'il n'y ait pas de cas de vandalisme, mais l'équipe en charge n'a rencontré aucun problème.

Le projet est considéré comme un succès relatif, le nombre de participants n'étant pas très élevé. Mme Regamey insiste cependant sur le fait que l'objectif premier était de faire découvrir la plateforme en elle-même, et que le fait que les usagers s'essayent ensuite à la correction d'OCR était plutôt perçu comme un plus. Les retours des usagers montrent qu'en général, ils étaient surpris en bien, contents de leur expérience, et regrettaient de ne pas avoir découvert la plateforme plus tôt. En cela, le projet est un succès, car ce sont 129 personnes de plus qui ont découvert e-newspaperarchives.ch et qui pourront en parler autour d'eux.

Mme Regamey affirme que ce projet était une très bonne expérience, qui n'était pas trop lourde à organiser, et qui en valait la peine. Elle recommande cependant de bien prendre garde à demeurer transparent, par exemple en faisant cocher une case, lors de l'inscription, indiquant si l'on accepte d'être contactés au sujet de nouveaux projets. Une collecte systématique des adresses e-mail des participants serait contraire au droit de la personnalité et leur utilisation trop fréquente pourrait agacer les usagers, ce qui serait dommageable pour la plateforme. Enfin, comme M. Glaser, elle recommande de « frapper fort », car le nombre d'usagers actifs redescend vite, et qu'il vaut donc mieux en recruter le plus possible dès le début.

4.2.4 Alexa Renggli – coordinatrice de la plateforme e-manuscripta.ch et responsable de projet pour l'outil de transcription – 04.06.2020

Alexa Renggli travaille pour la Zentralbibliothek Zürich et est responsable de l'outil de transcription collaborative sur e-manuscripta.ch, un projet pionnier en Suisse.

Mme Renggli explique que l'outil de transcription de la plateforme a été développé en interne avec l'entreprise employée pour créer la plateforme, Semantics Kommunikationsmanagement GmbH. Il s'agissait d'un besoin exprimé par toutes les institutions participantes, qui souhaitaient permettre la transcription des documents numérisés disponibles.

Sur cette plateforme, l'inscription est obligatoire et l'on doit indiquer son identité, ce qui permet déjà de limiter les « trolls », et simplifie la modération. Les transcriptions produites sont indexées et liées au catalogue, permettant ainsi la recherche plein texte. Il est possible de formater le texte avec des balises en Markdown, et les règles de transcription sont relativement simples. Mme Renggli souligne en effet qu'il lui semblait essentiel de proposer un standard minimum, valide pour tout type de manuscrit, et non d'imposer des règles trop spécifiques qui

pourraient compliquer la tâche des usagers et les décourager. La relecture des transcriptions est effectuée par les institutions participantes, sans être trop stricte. Il est toujours possible pour les usagers d'éditer et corriger les transcriptions par la suite, car un historique des versions permet un retour en arrière en cas de problème.

Pour la suite, il est prévu de continuer à améliorer l'outil en proposant la possibilité de tagger les documents, et de travailler à l'archivage à long terme des métadonnées produites dans le cadre de ce projet.

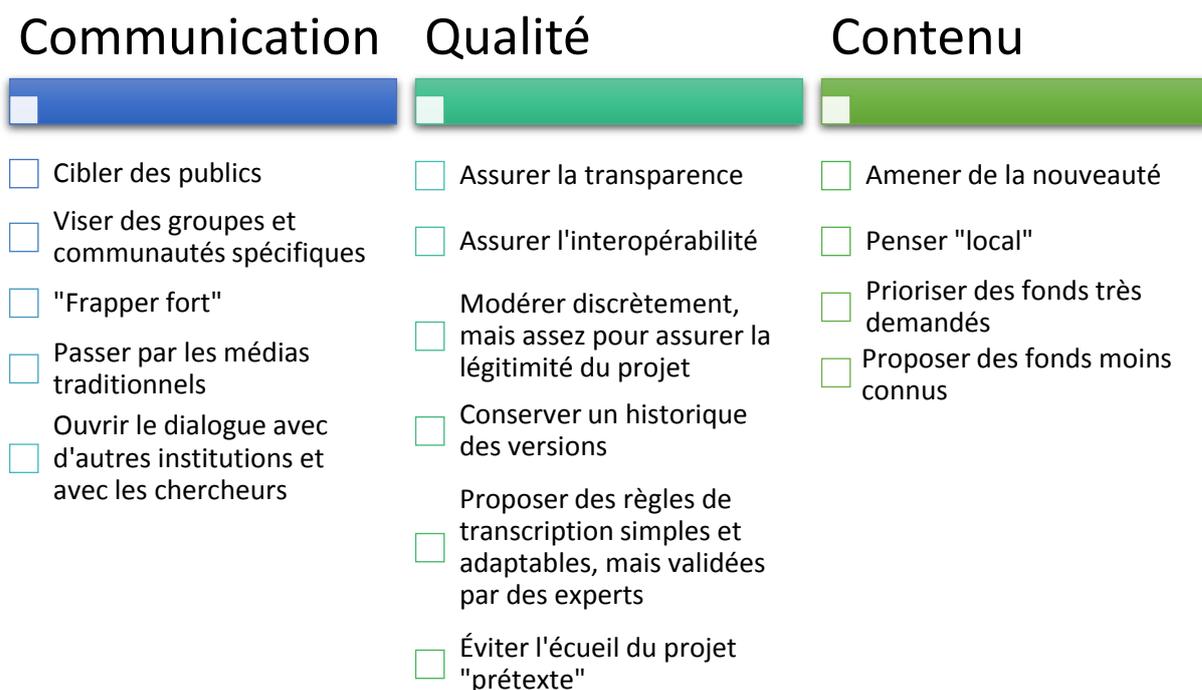
Mme Renggli note cependant que ce projet s'est construit sur demande des institutions, sans réellement définir un public cible. Leur but est de faire venir autant de monde que possible, mais de fait, il lui semble important de réfléchir à la question des publics pour adapter l'outil aux besoins des usagers. Elle mentionne notamment la possibilité d'utiliser la transcription collaborative dans le cadre de séminaires universitaires.

Concernant la perspective d'un tel projet en Ville de Genève, Mme Renggli pense que cela peut être bénéfique pour les institutions et les chercheurs. Elle insiste sur l'importance de l'interopérabilité pour enrichir les interfaces et mentionne qu'il peut être utile d'avoir recours aux protocoles [IIIF](#), comme le fait par exemple l'outil *FromThePage*. Enfin, elle rappelle la nécessité d'être transparent envers les usagers concernant leurs données et ce qui en est fait.

4.2.5 Éléments à retenir des entretiens

Plusieurs points importants, résumés dans la figure ci-dessous, ressortent des entretiens. En plus de ceux-ci, on retiendra qu'un projet de ce type peut être un succès sans pour autant mobiliser une foule immense. Au contraire, comme le signale Mme Regamey, toute nouvelle personne découvrant la plateforme est un succès en soi. Si l'objectif premier est de faire découvrir des collections et de permettre au public de se les approprier, il faut viser grand, mais voir chaque nouvelle inscription ou contribution comme une réussite.

Figure 22 : Éléments à retenir des entretiens avec les experts



4.3 Questionnaires et tests – DCTN et Comité éditorial

Les questionnaires et la grille d'expérience ont été envoyés aux dix-huit membres du Comité éditorial, ainsi qu'à une représentante du DCTN, avec un délai de réponse de deux semaines. Sept questionnaires (dont 2 du même répondant) et trois grilles d'expérience ont été retournés.

Une synthèse analytique des réponses a été envoyée aux personnes contactées, à laquelle étaient joints deux schémas créés sur la base des recherches et des réponses reçues. Cette partie en présente les principaux éléments.

4.3.1 Objectifs, attentes et besoins

Des réponses reçues ressortent les problématiques suivantes :

- **Visibilité des institutions, de leurs collections et de la plateforme** → se faire connaître
- **Lien entre les institutions** → collaborer, partager ses collections et ses compétences
- **Lien avec les publics** → collaborer, échanger, fidéliser
- **Participation citoyenne** → donner une place aux usagers, un rôle, des possibilités d'action, un pouvoir décisionnel, justifier l'utilisation des deniers publics
- **Démocratisation des savoirs** → assurer la transmission et la diffusion des savoirs
- **Accès aux collections** → faciliter la recherche et la découverte
- **Soutien au travail bibliothéconomique et archivistique** → conserver, compléter et enrichir les inventaires et les bases de données, diversifier les activités

De cela découlent les propositions suivantes :

Un espace de transcription collaborative devra...

- **Mettre en avant** les institutions, leurs collections et leurs compétences
- **Laisser une place** à chaque institution tout en encourageant les **collaborations entre institutions**
- Permettre de **communiquer** avec les usagers
- Être ouvert à la **participation** de tout un chacun, sans connaissance ou compétence spécifique préalable, et permettre aux usagers de **communiquer** entre eux
- Permettre aux usagers de continuer à **s'instruire**
- Permettre la création de **mots-clés** et la **recherche plein texte** dans les transcriptions
- Permettre de lier les données créées – et vérifiées – aux **catalogues** et autres **bases de données** existantes

Ces éléments seront au cœur du projet, car ils représentent les objectifs, besoins et attentes des institutions. La manière de les mettre en œuvre devrait *a priori* être discutée avec les usagers potentiels (cf. [publics](#), [usagers potentiels](#)), dans un esprit de participation.

4.3.2 Publics, usagers potentiels

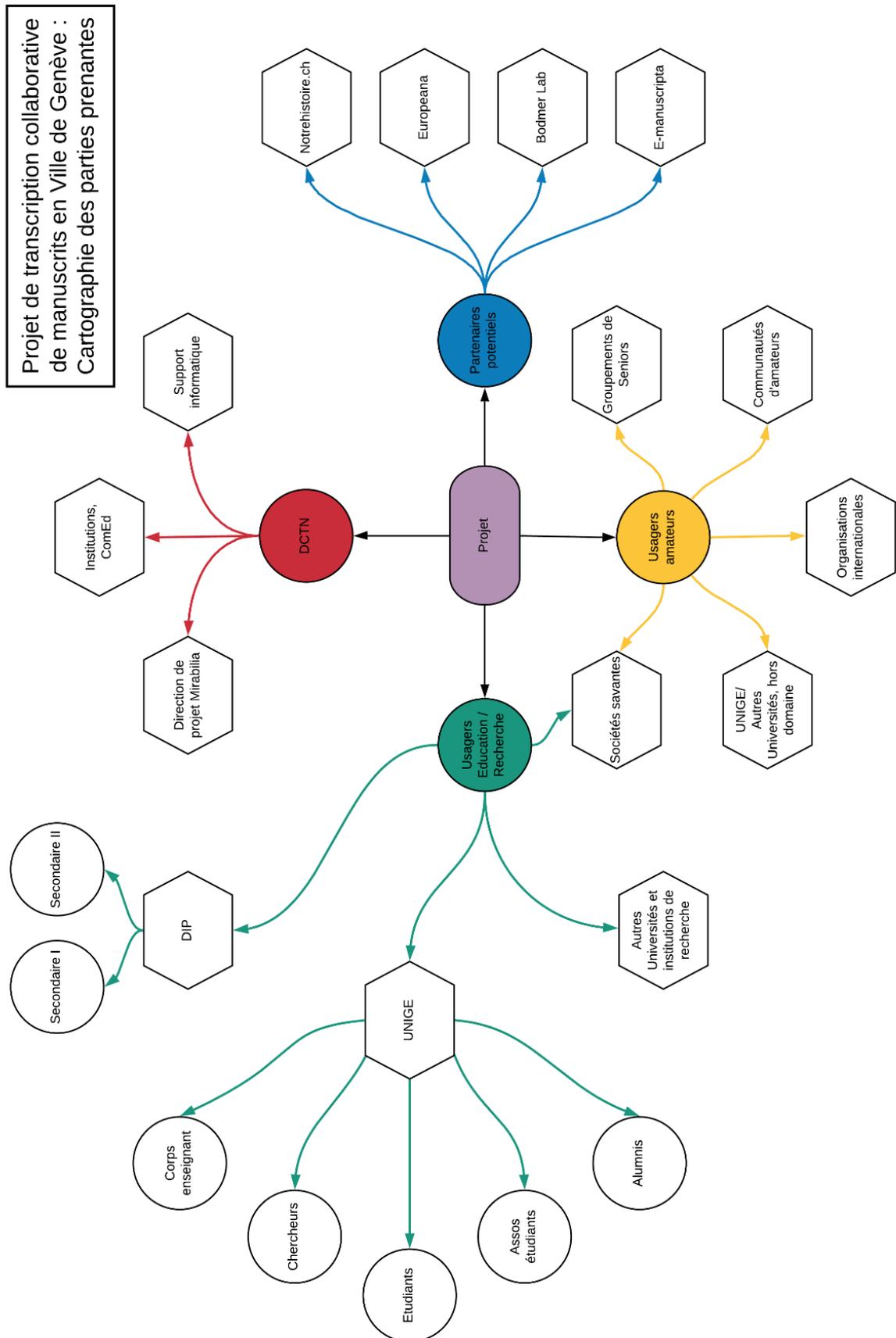
Les publics identifiés dans les questionnaires sont les suivants :

Tableau 1 : Publics potentiels identifiés par les tenants du projet

Publics cibles	Publics potentiels	Usagers extrêmes
Retraités (passionnés ou amateurs)	Collaborateurs d'autres institutions similaires	Retraités (amateurs)
Généalogistes	Généalogistes	Prisonniers
Scolaires	Scolaires	Gamers
Étudiants	Étudiants	Large public
Chercheurs	Chercheurs	
Amis des musées	Passionnés du domaine	
Habitants de Genève	Amateurs du domaine	
Communauté d'origine des documents	Usagers d'autres plateformes collaboratives	
	Participants à Wikipédia	

Certains publics se retrouvent dans différentes catégories, et il est évident que, selon les documents choisis, leurs thèmes, leurs auteurs ou leur contenu, différentes communautés de professionnels ou d'amateurs seront concernées. La communication devra donc s'y adapter. La [cartographie des parties prenantes](#) suivante doit permettre de simplifier ce processus. À noter que les associations et sociétés sont proposées comme points d'entrée vers les types de publics visés.

Figure 23 : Parties prenantes à prendre en compte dans un projet de *crowdsourcing* en Ville de Genève. Pour chaque projet, les parties prenantes doivent être précisées selon le(s) domaine(s) concerné(s) (par exemple « étudiants en Biologie », « chercheurs en Histoire de l'art » etc.)



4.3.3 Un projet participatif

Seule une institution mentionne l'expérience de projets participatifs, avec Wikipedia, outil basé sur la participation et pour lequel l'apport de communautés d'experts est essentiel. D'autres projets mentionnés concernent du « tagging » d'œuvres, ainsi que la « grande collecte » d'objets et documents de la Grande Guerre.

Si peu de projets de ce type ont fait parler d'eux à Genève, le participatif est dans l'air du temps, dans les bibliothèques (Bats 2015) comme dans les musées (ICOM 2019). Permettre aux citoyens d'être impliqués dans des projets d'institutions publiques est en effet vu comme une nouvelle manière de penser et de mettre en œuvre la démocratie. Mirabilia étant une plateforme à vocation participative, il serait cohérent que les publics soient également impliqués dans la gestion de l'espace de transcription collaborative.

Les niveaux de participation retenus sont les suivants :

- *Crowdsourcing* « plus libre », où les usagers peuvent contribuer de diverses manières
- Consultation, où ils sont en outre consultés sur certains éléments relatifs au projet

La « Conciliation », où les usagers et le Comité éditorial prennent les décisions en commun, est aussi envisageable, à condition qu'il s'agisse de communautés spécifiques, dont l'expertise serait utile, ou d'« habitués », une fois que la plateforme aura trouvé son public.

Dans un premier temps, il est envisageable de viser le *crowdsourcing* « plus libre ». Les groupes d'usagers potentiels formés pourront cependant être consultés sur certains aspects.

Par la suite, en fonction des résultats et de la fréquentation de la plateforme, on pourra offrir plus de possibilités de participation, par exemple en formant des comités d'usagers habitués de la plateforme et/ou de spécialistes (professionnels ou amateurs).

4.3.4 Sélection des documents

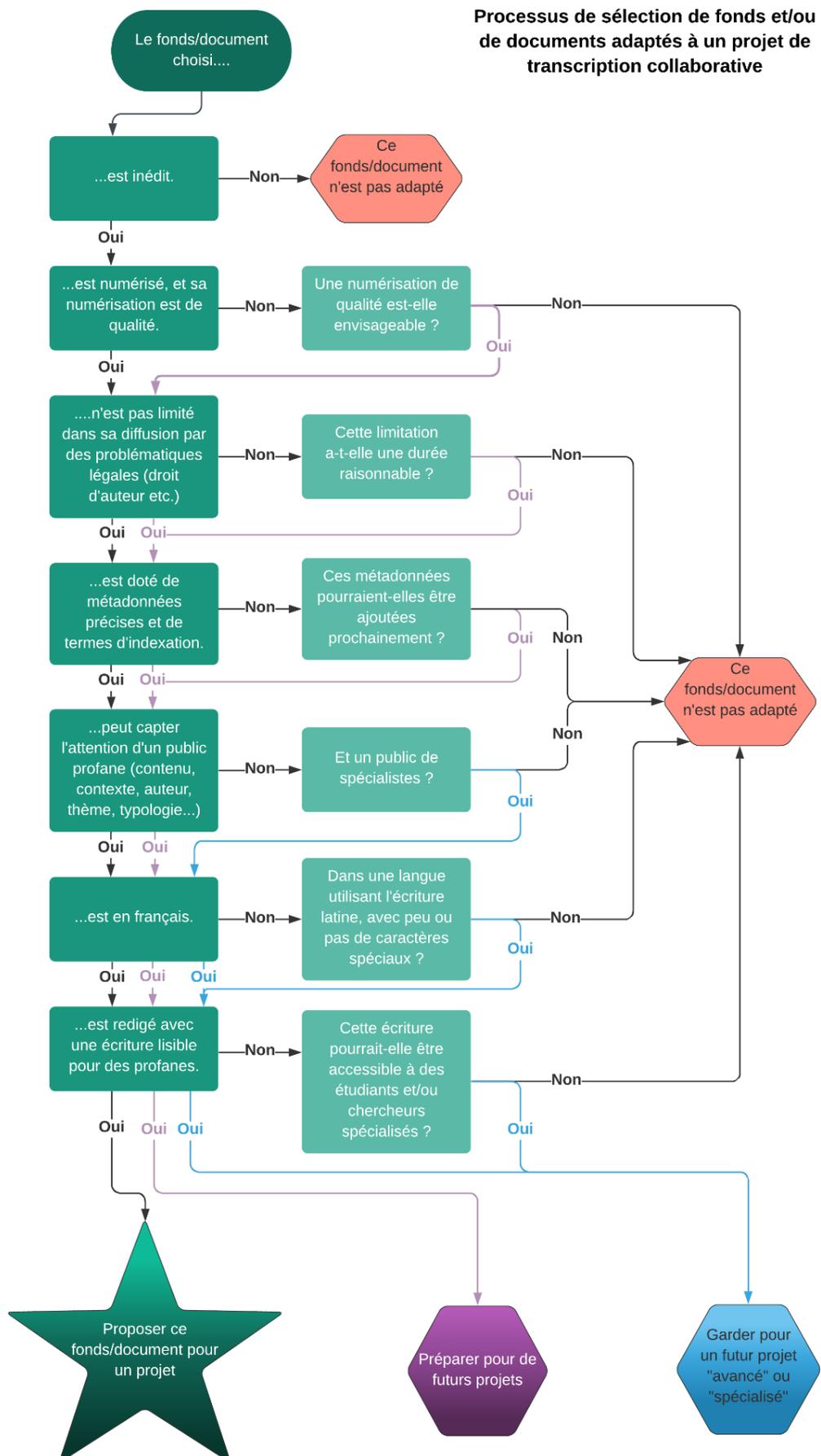
Les documents proposés sont très variés (journaux satiriques, récits de voyage, collection de sceaux, registres d'État civil et des cimetières, correspondances, étiquettes de cadres entomologiques etc.), ce qui présage d'une belle diversité de futurs projets.

Les [entretiens](#) effectués avec des personnes ayant mené des projets de *crowdsourcing* en Suisse montrent qu'il faut qu'un premier projet « frappe fort », qu'il touche à des questions d'actualité, et qu'il attire l'œil d'une grande variété de personnes. Il faudra donc choisir en ce sens. Par la suite d'autres projets moins « frappants », visant des communautés plus spécifiques, pourront être menés.

L'important reste que chaque institution puisse trouver sa place avec des projets de *crowdsourcing* qui lui sont propres, mais que les institutions collaborent également, avec par exemple des projets les rassemblant autour d'une même thématique

La question de qui choisit les documents reste complexe. Il va de soi que l'expertise des responsables de collections restera au cœur de chaque projet, et le [guide décisionnel](#) créé à cet effet devrait faciliter cette sélection. Il serait cependant envisageable, dans un esprit de participation, de présenter aux usagers une présélection de fonds identifiés par les institutions, parmi lesquels choisir, voire les laisser proposer eux-mêmes des fonds.

Figure 24 : Guide décisionnel pour la sélection de fonds



4.3.5 Plateformes testées

Les répondants ont testé Europeana, e-manuscripta et Library of Congress : By the people.

Les principaux points relevés sont les suivants :

- Avoir une **page d'accueil claire**, pour que l'on sache où cliquer
- Mettre l'accent sur le **visuel**, qui attire le regard
- Mettre à disposition des **tutoriels** (vidéo, idéalement) et **modes d'emploi** clairs
- Mettre à disposition, et à portée de clic, les **règles de transcription**
- **Varié** les types de documents, mais les présenter dans un **ordre logique**
- Rendre visible la **progression** de la transcription de chaque document
- Clarifier le système de **révision** et **validation**
- Accompagner les documents de **commentaires ou billets de spécialistes** permettant de remettre les œuvres dans leur **contexte**
- Proposer des liens vers des **ressources externes**
- **Lier les données créées** à des bases de données existantes, voire à Wikipédia
- Essayer de travailler directement avec une **plateforme existante**, plutôt que d'en créer une de toute pièce

Dans l'ensemble, ces constats rejoignent ceux des [observations participantes](#). Il est cependant nécessaire de les comparer aux [retours des usagers potentiels](#).

4.4 *Persona* et recrutement d'usagers potentiels

Le [tableau](#) des publics potentiels identifiés par les représentants du DCTN et le Comité éditorial, ainsi que les cartes d'empathie réalisées par les répondants, ont servi de base à la réalisation de *persona* d'usagers potentiels. Certains profils pouvant être combinés, tous n'apparaissent pas. Le large public, ou l'habitant de Genève peut en effet correspondre à plusieurs des profils proposés. L'ami des musées peut être assimilé à l'amateur, au sens où l'entend [Trevor Owens](#), et les communautés d'origine peuvent être considérés comme des chercheurs ou à des amateurs, selon leurs objectifs. Cette dernière catégorie est plus particulièrement liée aux fonds du MEG, et est donc moins généralisable, d'où son absence. Enfin, les prisonniers ont été remplacés par la *persona* « Bibliothécaire de prison », figure plus accessible pour discuter de la disponibilité d'un tel projet dans ce type d'établissement.

La figure ci-dessous présente des exemples de *persona*. La totalité d'entre elles se trouve en [annexe](#). La couleur d'encadrement permet de les distinguer lors d'une prochaine étape du travail, les [parcours utilisateurs](#).

Figure 25 : Exemples de *persona*

<p>Profil A : Personne à la retraite</p>  <p>Nom : Isabelle A. Âge : 68 ans Activité : Retraitée Centres d'intérêt : Art, littérature, culture, histoire Appareils à disposition : ordinateur portable User story : « Étant à la retraite, je fréquente régulièrement les institutions culturelles genevoises, et je souhaiterais consacrer une partie de mon temps à me rendre utile à ma communauté, tout en m'instruisant et me divertissant. »</p> <p>Lien de l'image</p>	<p>Profil C : Enseignant universitaire</p>  <p>Nom : Stéphane C. Âge : 55 ans Activité : Chargé d'enseignement en Histoire moderne Centres d'intérêt : Histoire de Genève, art moderne, tennis Appareils à disposition : ordinateur portable et de bureau, smartphone User story : « J'aime pouvoir aider des amateurs d'Histoire à aborder des sources et à en comprendre le contenu et l'importance. En outre, j'aimerais que mes étudiants puissent s'exercer à la paléographie, afin d'être autonome dans leurs recherches et éditions de textes. »</p> <p>Lien de l'image</p>
<p>Profil B : Amateur d'histoire, culture, littérature</p>  <p>Nom : Sylvain B. Âge : 38 ans Activité : Community Manager Centres d'intérêt : Photographie, littérature, histoire, paddle Appareils à disposition : ordinateur portable, smartphone User story : « Je n'ai pas beaucoup de temps à cause de mon travail et de ma vie de famille, mais j'aime bien l'Histoire et la littérature, et si je peux me rendre utile en m'instruisant de temps en temps je suis ravi. »</p> <p>Lien de l'image</p>	<p>Profil D : Enseignant au secondaire I ou II</p>  <p>Nom : Camille D. Âge : 32 ans Activité : Enseignante d'Histoire au cycle et au collège Centres d'intérêt : Histoire, sciences, danse moderne Appareils à disposition : ordinateur portable, smartphone User story : « Je cherche toujours de nouvelles idées pour créer des séquences pédagogiques innovantes et interactives. De plus, je trouve essentiel que les élèves fassent l'expérience d'être confrontés à des sources brutes, pour mieux comprendre le travail des historiens. »</p> <p>Lien de l'image</p>
<p>Profil E : Généalogiste</p>  <p>Nom : Philippe E. Âge : 42 ans Activité : Comptable Centres d'intérêt : Généalogie, histoire locale, littérature suisse Appareils à disposition : ordinateur portable, smartphone User story : « Je consacre presque tout mon temps libre à des recherches en généalogie, et un accès direct aux documents qui me sont utiles, en ligne depuis chez moi, m'aiderait beaucoup dans mon travail, surtout s'ils sont déjà transcrits. »</p> <p>Lien de l'image</p>	

Comme signalé [plus haut](#), le recrutement des usagers potentiels a commencé en amont de la réalisation des *persona*. Dans le cas des amateurs, une distinction par tranche d'âge a été effectuée, afin de mieux représenter la variété des personnes concernées. Les groupes d'usagers potentiels suivants ont ainsi été formés :

- Enseignement et recherche en Histoire
 - Enseignant/chercheur universitaire – actif sur une plateforme de *crowdsourcing*
 - Enseignant/chercheur universitaire – membre d'une société savante
 - Etudiant/chercheur universitaire – membre d'une association d'étudiants
 - Enseignant au secondaire récemment retraité – membre d'une société savante
 - Chercheur amateur
- Amateurs
 - Personne à la retraite, intéressée par l'Art, la littérature et l'Histoire, fréquentant régulièrement les institutions culturelles de la Ville de Genève
 - Personne en activité proche de la retraite, Alumni d'UNIGE diplômé en Histoire, fréquentant régulièrement les institutions culturelles de la Ville de Genève
 - Personne en activité (plus de 35 ans), intéressée par l'Art et l'Histoire

- Personne en activité (moins de 35 ans), intéressée par les Sciences et la culture en général, mais fréquentant peu les institutions culturelles genevoises

Ont été également contactés, sans succès, la Société de Lecture, les Alumni de l'UNIGE, des enseignants du secondaire I et II en activité, l'équipe de didactique de l'Histoire et de la citoyenneté de l'Institut universitaire de Formation des enseignants (IUFÉ), l'Université du 3e âge et Cité Seniors. De ce fait, et en raison des limites de mon propre réseau et du manque de temps, certains profils n'ont pas pu être représentés.

À la suite de la réalisation des *persona*, des ajustements ont pu être opérés, et deux autres catégories d'utilisateurs potentiels ont été créées :

- Généalogistes amateurs
 - Personne en activité, membre de la Société Genevoise de Généalogie
 - Personne à la retraite, membre de la Société Genevoise de Généalogie
- Enseignement et recherche en Sciences naturelles
 - Personne récemment diplômée en Biologie et active dans ce domaine

Ces quatre groupes ne représentent pas la totalité du spectre des possibles, mais ils ont permis de recevoir des retours divers et de commencer à faire connaître le projet à des utilisateurs potentiels, mais aussi à des utilisateurs finaux des transcriptions produites – les chercheurs et généalogistes. Ces utilisateurs potentiels sont en effet essentiels, non seulement parce que, du fait de la méthode choisie, leur avis et suggestions doivent être pris en compte à chaque étape, mais aussi car, si le projet les convainc et leur plaît, ils en seront les premiers ambassadeurs.

4.5 Questionnaires et tests – utilisateurs potentiels

À chaque groupe d'utilisateurs, une synthèse analytique des résultats a été transmise, généralement par oral en visioconférence, et cette partie en présente la teneur.

4.5.1 Enseignement et recherche en Histoire

Pour ce groupe, la grille et le questionnaire ont été transmis à cinq personnes, avec un délai de retour de deux semaines, mais l'une des cinq personnes, à savoir celle qui est active sur une autre plateforme de *crowdsourcing*, s'est retirée pour des raisons personnelles. Quatre questionnaires et cinq grilles d'expérience, dont deux de la même personne, ont été retournés.

4.5.1.1 Plateformes testées

Les répondants ont testé les plateformes Europeana, e-manuscripta et TACT.

Les principaux points relevés sont les suivants :

- Présenter clairement **le projet et ses enjeux**, le **contexte** des documents etc.
- Expliciter les **consignes** et le système de **création d'un compte**
- Proposer un **corpus délimité**, **classé** logiquement et proposant des **fonctions de recherches** et de **filtres**

- Lorsqu'un document comporte plusieurs pages, le présenter dans son **intégralité** avec la possibilité de passer à la page suivante, et pas un folio à la fois, en évitant les doubles pages
- Proposer un **tutoriel** et des **règles de transcription** claires, et à portée de clic
- Assurer une **bonne qualité de l'image numérisée**, pour pouvoir zoomer
- Donner la possibilité de **retrouver les documents que l'on a transcrit** sur son compte
- Clarifier le **système de validation**
- Assurer la **transparence** quant à la protection des données

4.5.1.2 Questionnaires

Pour chacune des quatre questions, les points principaux sont relevés ci-dessous.

Intérêts perçus pour un projet de transcription collaborative en Ville de Genève :

- Enseignement supérieur
 - Exercices de paléographie
 - Travail avec des sources lors de séminaires et de travaux de Bachelor
 - Édition de textes lors de travaux de Master
 - Accès aux documents facilité par les transcriptions et indexations produites
- Enseignement secondaire
 - Intégration à une séquence pédagogique et exploitation du contenu transcrit
 - Présentation de certains fonds méconnus
 - Utilisation lors de travaux de Maturité
- Recherche
 - Gain de temps pour l'édition et la recherche dans les catalogues
 - Mutualisation des chercheurs, si l'on peut créer des projets à accès limité
 - Accès direct aux documents numérisés
- Loisirs
 - Généalogie
 - S'instruire en s'amusant
 - Se sentir utile sur son temps libre

L'un des répondants insiste sur la nécessité du dialogue avec les enseignants et chercheurs et de la prise en compte du tissu associatif genevois pour promouvoir et donner sens à ce projet.

Besoins auxquels pourrait répondre ce projet :

- Accès aux numérisations
- Simplification du repérage et de l'identification des sources
- Élargissement du champ de recherche documentaire

Certains répondants ne relèvent aucun besoin, mais pensent que la plateforme peut en créer.

Attentes concernant le projet :

- Meilleure accessibilité des documents grâce aux numérisations et aux métadonnées
- Meilleure compréhension du rôle du patrimoine par la population genevoise
- Découverte des personnalités genevoises et de l'Histoire locale, à travers des thèmes comme les guerres mondiales, la crise des années 1930, la Société des Nations etc.

Craintes des répondants :

- Nombre d'erreurs et temps de correction
- Définition du corpus de documents, qui doit être intéressant et utile pour les usagers
- Risque que ce type de plateforme finisse par remplacer l'édition scientifique réalisée par des professionnels

De ces réponses découle que les enseignants et chercheurs seraient plutôt prêts à être des utilisateurs finaux des transcriptions produites que des usagers de la plateforme elle-même. Ils pourraient jouer ce second rôle à condition que la plateforme s'adapte à leurs besoins, en leur offrant la possibilité, par exemple, de demander la mise à disposition de certains documents en vue d'un enseignement ou d'un projet de recherche.

Le projet aurait alors un double objectif pour eux : mettre à disposition les documents numérisés selon les besoins, et jouer le jeu de la transcription collaborative pour produire de nouvelles métadonnées.

Enfin, en réponse aux craintes exprimées, il est certain que les tutoriels, règles de transcription et circuits de validation devront être bien élaborés pour limiter les erreurs. Pour le corpus, intégrer les usagers à sa définition, comme suggéré [plus haut](#), semble une bonne solution. Enfin, il est aisé de répondre à la troisième crainte, car l'expertise des spécialistes demeure nécessaire dans toute édition scientifique, même celles qui impliquent du *crowdsourcing*, comme le montre le projet Transcribe Bentham.

4.5.2 Amateurs

Pour ce groupe, la grille et le questionnaire ont été transmis à quatre personnes, avec un délai de retour de deux semaines. Trois grilles et trois questionnaires ont été retournés.

4.5.2.1 Plateformes testées

Les répondants ont testé les plateformes e-manuscripta et TACT.

Les principaux points relevés sont les suivants :

- Commencer par une **page d'accueil** claire, sobre et esthétique, qui présente le **contexte** du projet, son **état d'avancement** etc.
- Proposer un **tutoriel** et des **consignes** simples mais précises
- Rassurer les usagers sur la **légitimité** de leurs savoirs et compétences
- Pouvoir affiner les transcriptions par du **balisage** pour mieux appréhender l'œuvre. Si cela semble **complexe** de prime abord, cela devient **ludique** par la suite.
- Sentir que l'on contribue, que **l'on se rend utile**, même si cela nécessite de la **rigueur**

Il est intéressant de noter ici que le système de balisage, considéré comme trop complexe lors des **observations participantes**, a été apprécié par les usagers amateurs, ce qui prouve l'importance de ce type de tests : les usagers se montrent souvent capables de bien plus que ce qu'on aurait pu croire.

4.5.2.2 Questionnaires

Ce questionnaire diffère légèrement du précédent. En effet, les « amateurs » ne côtoyant pas les institutions culturelles de la Ville de Genève dans le cadre professionnel⁶, il est intéressant de les interroger sur leur fréquentation de ces institutions et leur attachement à elles.

Les répondants connaissent toutes les institutions concernées ou presque, et en fréquentent quelques-unes au moins. Ils les apprécient pour la richesse de leurs collections et l'ambiance qui s'y dégage, et s'y rendent avant tout pour apprendre et se divertir. Certains mentionnent un attachement plus personnel, lié à des souvenirs d'enfance ou à leur carrière. Leur image des institutions est donc positive, et il s'agit d'un public disposé à tester et découvrir de nouveaux services mis à disposition par ces dernières.

La suite du questionnaire reprend la forme du précédent, avec une question supplémentaire :

Intérêts perçus pour un projet de transcription collaborative en Ville de Genève :

- Participer à une meilleure connaissance du patrimoine et à le rendre plus accessible
- S'instruire sur son temps libre
- Rencontrer d'autres personnes intéressées par les mêmes sujets
- Mieux appréhender certaines œuvres, certains auteurs et leur contexte

Besoins auxquels pourrait répondre ce projet :

- Améliorer les recherches, l'accès aux documents et à l'information
- Faire découvrir et rendre accessible des documents patrimoniaux

Attentes concernant le projet :

- Créer des communautés d'intérêt

⁶ À l'exception d'un participant, encore employé par une institution genevoise au moment de répondre au questionnaire.

- Contribuer à la recherche, être utile aux chercheurs et étudiants
- Rendre le projet accessible à tous, de manière ludique
- Assurer la transparence du projet
- Communiquer suffisamment pour attirer les gens

Craintes des répondants :

- Choix d'ouvrages trop spécialisés et communautés trop fermées
- Outil difficile à prendre en main
- Sentiment d'illégitimité
- Manque d'intérêt de la part du public
- Communication trop confidentielle

Types de documents, **thèmes** et **auteurs** qui pourraient les intéresser :

- Documents des XIXème et XXème siècles
- Manuscrits d'auteurs
- Correspondance scientifique
- Littérature, art, histoire, biographie, poésie
- Sciences humaines et sociales, ethnographie, anthropologie
- Sciences naturelles
- Thèmes morbides (*sic*), procès de sorcellerie, cimetières etc.
- Auteurs genevois
- Auteurs francophones et anglophones contemporains

La dernière question n'a été posée qu'à ce groupe, car les autres auraient pu avoir des réponses trop spécifiquement liées à leurs champs de recherche ou d'enseignement. Il était intéressant, en revanche, de comprendre ce qui pouvait capter l'attention des amateurs, et leurs réponses montrent qu'une large variété de thématiques peut être abordée.

Les réponses ci-dessus dénotent un certain intérêt pour un projet de ce type. La notion de communauté et de rencontre revient à plusieurs reprises, et est perçue comme une plus-value.

La majorité des craintes tournent autour du risque d'un clivage élitiste entre spécialistes et non-spécialistes, de ne pas se sentir accueilli, utile et légitime. L'ergonomie et les outils didactiques de la plateforme, ainsi que la communication et la médiation devront veiller à pallier ce risque. Enfin, pour éviter le désintérêt du public, être à son écoute et l'intégrer aux discussions demeure la meilleure solution.

4.5.3 Généalogie

Pour ce groupe, la grille et le questionnaire ont été transmis à deux personnes, avec un délai de retour de deux semaines. Une grille et un questionnaire ont été retournés.

4.5.3.1 Plateforme testée

Le répondant a testé la plateforme e-manuscripta.

Il a surtout apprécié les points suivants :

- Proposer un **outil intuitif** et une bonne **ergonomie**
- Permettre de choisir la **langue**, si la plateforme est multilingue
- Proposer une **diversité** de documents, mais faire attention aux **degrés de difficulté**

4.5.3.2 Questionnaire

Pour chacune des quatre questions, les points principaux sont relevés ci-dessous.

Intérêts perçus pour un projet de transcription collaborative en Ville de Genève :

- Généalogie
 - Accès facilité aux documents complexes à déchiffrer
 - Recherche de noms grâce à l'indexation et à des mots-clés
- Autres domaines
 - Recherche historique
 - Initiation à la paléographie

Besoins auxquels pourrait répondre ce projet :

- Compléter les documents disponibles sur la plateforme de la Société Genevoise de Généalogie avec un relevé des registres (mariages, décès) disponibles.

Attentes concernant le projet :

- Pouvoir travailler sur des documents manuscrits genevois (registres, minutes etc.)

Craintes du répondant :

- Limites liées aux fonds disponibles, si le projet exclut les Archives d'État de Genève

Les généalogistes sont un public cible de premier ordre pour un tel projet. La crainte exprimée pose cependant un problème car, s'agissant d'un projet « Ville de Genève », les AEG devraient *a priori* être exclues. Il conviendrait donc d'identifier des documents au sein des AVG qui pourraient répondre aux besoins de ces usagers, tels que les registres des cimetières.

4.5.4 Enseignement et recherche en Sciences naturelles

La grille et le questionnaire ont été transmis à la personne concernée, qui a retourné dans un délai de deux semaines le questionnaire et deux grilles d'expérience.

4.5.4.1 Plateformes testées

Le répondant a testé les plateformes e-manuscripta et les Herbonautes, qui lui avait été spécialement proposée du fait de son profil.

Les principaux points relevés sont les suivants :

- Créer un **tutoriel** obligatoire et intuitif
- Proposer un **choix aléatoire** de document, pour pouvoir commencer rapidement

Il est intéressant de noter que ce répondant est le seul à avoir exprimé le souhait de pouvoir faire un choix aléatoire. Cela peut être dû au type de document, car un herbier ne se lit pas nécessairement dans un ordre précis, à l'inverse d'un roman, par exemple.

4.5.4.2 Questionnaire

Intérêts perçus pour un projet de transcription collaborative en Ville de Genève :

- Enseignement
 - Aucun, *a priori*
- Recherche
 - Soutenir les scientifiques, qui n'ont ni le temps ni les moyens de tout transcrire
- Loisirs
 - Se prendre au jeu de transcrire des étiquettes descriptives de faune et de flore
 - Avoir accès à des ressources fragiles et précieuses
 - Découvrir la flore suisse
 - Contribuer à la science, permettre de découvrir des espèces menacées, de mieux comprendre la répartition géographique et l'évolution des espèces

Besoins auxquels pourrait répondre ce projet :

- Effectuer des recherches dans des herbiers numérisés

Attentes concernant le projet :

- Présenter la flore locale
- Proposer des filtres de classification, s'il s'agit de reconnaissance de faune et de flore
- Proposer une interface intuitive permettant de débiter rapidement, avec un tutoriel

Craintes du répondant :

- Méconnaissance de la botanique pouvant cause des erreurs

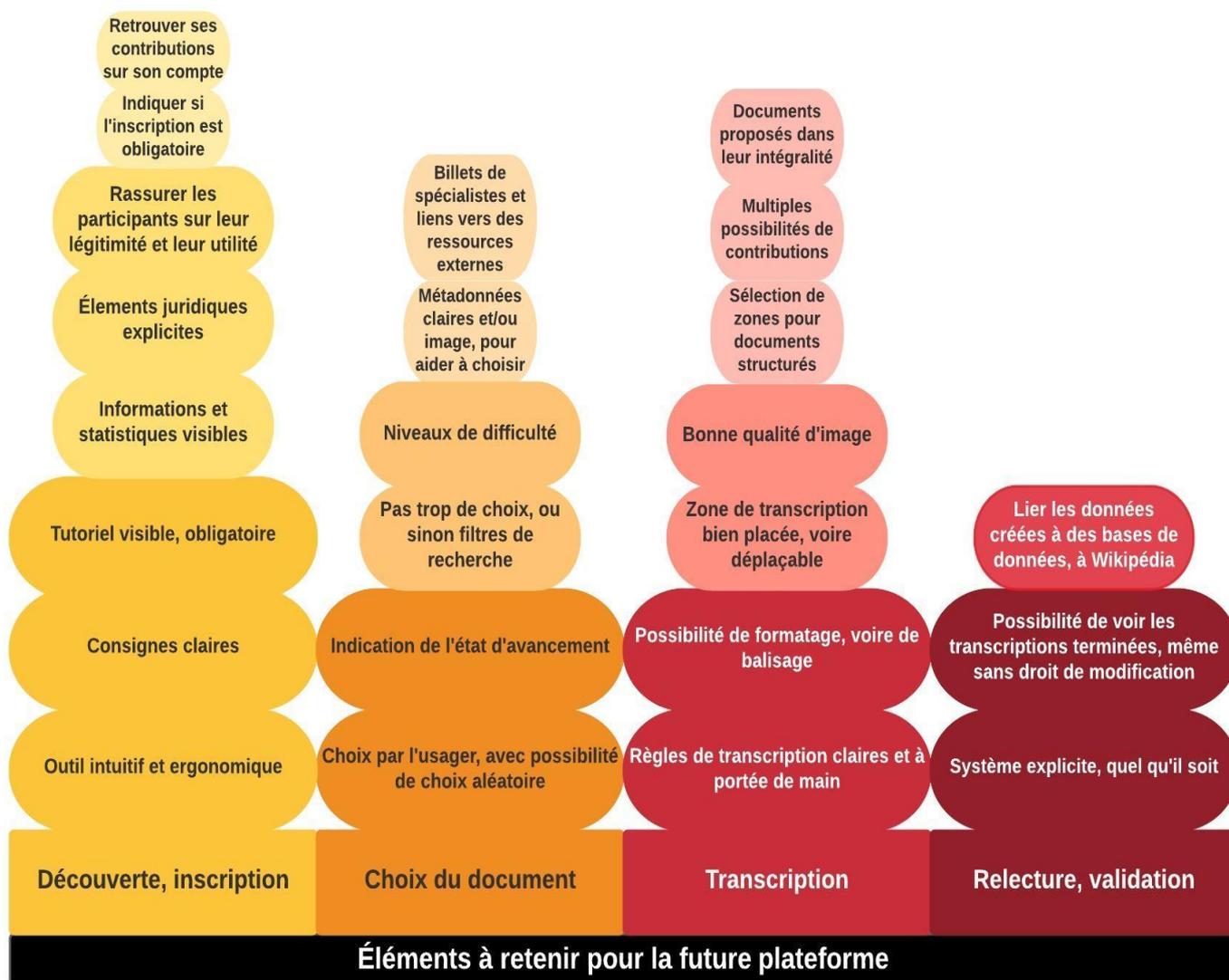
Comme pour le premier groupe, l'utilité de la plateforme résiderait dans l'accès aux documents permis par les métadonnées créées. Ce répondant semble tout de même trouver intéressant de contribuer sur son temps libre, pour pouvoir faciliter son travail et celui d'autres chercheurs.

Sa crainte rejoint celle du premier groupe : que des non-spécialistes multiplient les erreurs. Le circuit de validation doit donc impérativement être efficace.

4.5.5 Récapitulation générale

Chacun des groupes interrogés a permis de dégager des idées, propositions et problématiques différentes, qui permettront de mieux cadrer le projet et feront l'objet de [recommandations](#). Sur cette base, et sur la base des retours du Comité éditorial et des représentants du DCTN, la [figure 21](#) a été mise à jour comme suit :

Figure 26 : Éléments à retenir pour la future plateforme – mis à jour



4.6 Tri de cartes

Après les questionnaires et grilles d'expériences, les usagers potentiels ont participé à un tri de cartes en ligne, réalisé grâce à la plateforme [Optimalworkshop](#). Le but de cette étape est de trier différents éléments relatifs au projet et au futur outil de transcription collaborative en fonction de l'intérêt et de l'importance perçus. Un élément peut en effet susciter de l'intérêt, de l'enthousiasme, mais ne pas être perçu comme absolument utile, et un autre peut être considéré comme utile mais peu captivant en soi. En outre, certains éléments peuvent être considérés comme dépendant des professionnels, chercheurs ou spécialistes de l'information.

Les catégories dans lesquels les cartes devaient être triées sont donc les suivantes :

- Très intéressant, important : à faire absolument
- Très intéressant, pas important
- Plutôt intéressant, important
- Plutôt intéressant, pas important
- Peu intéressant, important
- Peu intéressant, pas important : à ne pas faire
- Seulement pour les spécialistes (bibliothécaires, archivistes et/ou chercheurs)

Les éléments à trier dans ces catégories ont été sélectionnés en fonction des différences entre les plateformes existantes, des retours des personnes en ayant testé et de discussions avec les responsables de Mirabilia. Ils étaient présentés aux participants dans un ordre aléatoire mais devaient tous être triés :

- Choisir son document librement
- Pouvoir revenir à sa transcription plus tard
- Vérifier et corriger les transcriptions des autres
- Signaler lorsque l'on a un doute sur la transcription d'un mot
- Poser des questions et discuter des textes transcrits avec des spécialistes
- Poser des questions et discuter des textes transcrits avec d'autres usagers
- Mettre un tag si un mot-clé, une thématique ressort du texte
- Signaler une localisation si un nom de lieu est mentionné dans le texte
- Signaler une localisation si l'on arrive à déterminer la provenance du document
- Signaler une date, si celle-ci est mentionnée dans le texte
- Signaler une date, si l'on arrive à définir celle du document lui-même
- Mettre un tag si un nom de personne est mentionné dans le texte

- Pouvoir voir si un document est déjà transcrit partiellement ou totalement avant de le sélectionner
- Ajouter ses propres documents manuscrits
- Être informé de l'avancée des transcriptions, par exemple lorsqu'on arrive à la moitié du fonds, ou vers la fin
- Être invité à un événement de clôture lorsque le fonds est transcrit
- Voir son nom inscrit comme contributeur dans l'édition du texte, si celui-ci fait ensuite l'objet d'une édition (papier ou numérique)
- Recevoir un exemplaire du texte édité, si celui-ci fait l'objet d'une édition (papier ou numérique)
- Recevoir un "badge", un trophée numérique, toutes les X lignes transcrites
- Avoir le droit de vérifier et corriger les transcriptions d'autrui, au bout de X lignes transcrites
- Être informé quand de nouveaux documents sont mis à disposition
- Être invité à des conférences concernant le fonds en cours de transcription
- Être informé de l'utilité des transcriptions pour les chercheurs, si des articles sont publiés sur le sujet, par exemple
- Être entièrement anonyme, participer sans avoir besoin de créer un compte
- Pouvoir donner mon nom complet, pour pouvoir être cité
- Transcrire librement, sans se préoccuper des règles de transcription
- Recevoir une notification lorsque l'on n'a pas transcrit depuis un certain temps
- Voir les transcriptions des autres, sans les corriger
- Pouvoir me mettre un "défi" de transcrire X lignes par jour
- Avoir accès à la plateforme depuis mon téléphone portable

Huit personnes ont participé à ce tri de cartes. La figure suivante présente les résultats de ce tri. Les textes sont malheureusement coupés, mais ils sont classés dans l'ordre ci-dessus.

Figure 27 : Résultats du tri de carte en ligne

	Très intéressant, ...	Très intéressant, ...	Plutôt intéressan...	Plutôt intéressan...	Peu intéressant, i...	Peu intéressant, ...	Seulement pour l...
Choisir son document librement	7		1				
Pouvoir revenir à sa transcription ...	8						
Vérifier et corriger les transcrip...	1		1		2		4
Signaler lorsque l'on a un doute s...	7		1				
Poser des questions et discuter d...	3		3	2			
Poser des questions et discuter d...	2		4	2			
Mettre un tag si un mot-clé, une t...	1	1	2	1			3
Signaler une localisation si un no...	5		1		1		1
Signaler une localisation si l'on arr...	5		1				2
Signaler une date, si celle-ci est m...	5			1		1	1
Signaler une date, si l'on arrivé e d...	5			1			2
Mettre un tag si un nom de perso...	4	1	1				2
Pouvoir voir si un document est d...	6		2				
Ajouter ses propres documents m...				1		3	4
Être informé de l'avancée des tra...	2	1	2	2			1
Être invité à un événement de clô...		2	2	1		2	1
Recevoir un exemplaire du texte é...		1	1	3		1	2
Voir son nom inscrit comme contri...	2	1	2	2		1	
Recevoir un "badge", un trophée ...		1		4		3	
Avoir le droit de vérifier et corrige...	1		3			1	3
Être informé quand de nouveaux ...		1	4	1			2
Être invité à des conférences con...			2	3	1		2
Être informé de l'utilité des transcr...	3	1	2	1			1
Être entièrement anonyme, partici...	1		1	1		5	
Pouvoir donner mon nom complet...	1			3	1	3	
Transcrire librement, sans se préo...			1	2		5	
Recevoir une notification lorsque l...	1			2		5	
Voir les transcriptions des autres, ...	5		2		1		
Pouvoir me mettre un "défi" de tra...		1		3		4	
Avoir accès à la plateforme depui...			1	3		4	

Ces résultats permettent de mieux comprendre comment les usagers potentiels perçoivent ces éléments, et de prioriser ceux qui ont été le plus plébiscités.

Outre les éléments qui ressortent comme « à faire absolument », on note le fait que des rencontres en présentiel, une gratification physique (cadeau, citation du nom dans un livre) et une forme de gamification, voire de compétition, n'intéressent que peu les répondants, ce qui rejoint les propos de [David Glaser](#).

On note également que les répondants sont divisés quant à la vérification et validation des transcriptions. La moitié d'entre eux considèrent que cette tâche revient aux spécialistes, alors que les autres envisagent qu'elle puisse être effectuée par des usagers, de préférence ceux avec une certaine expérience, comme c'est le cas sur Europeana, par exemple. Le choix de ce système dépend de la stratégie et des objectifs du projet, mais il est intéressant de prendre cette division en compte, pour pouvoir communiquer à ce sujet de manière adéquate.

4.7 Parcours utilisateurs

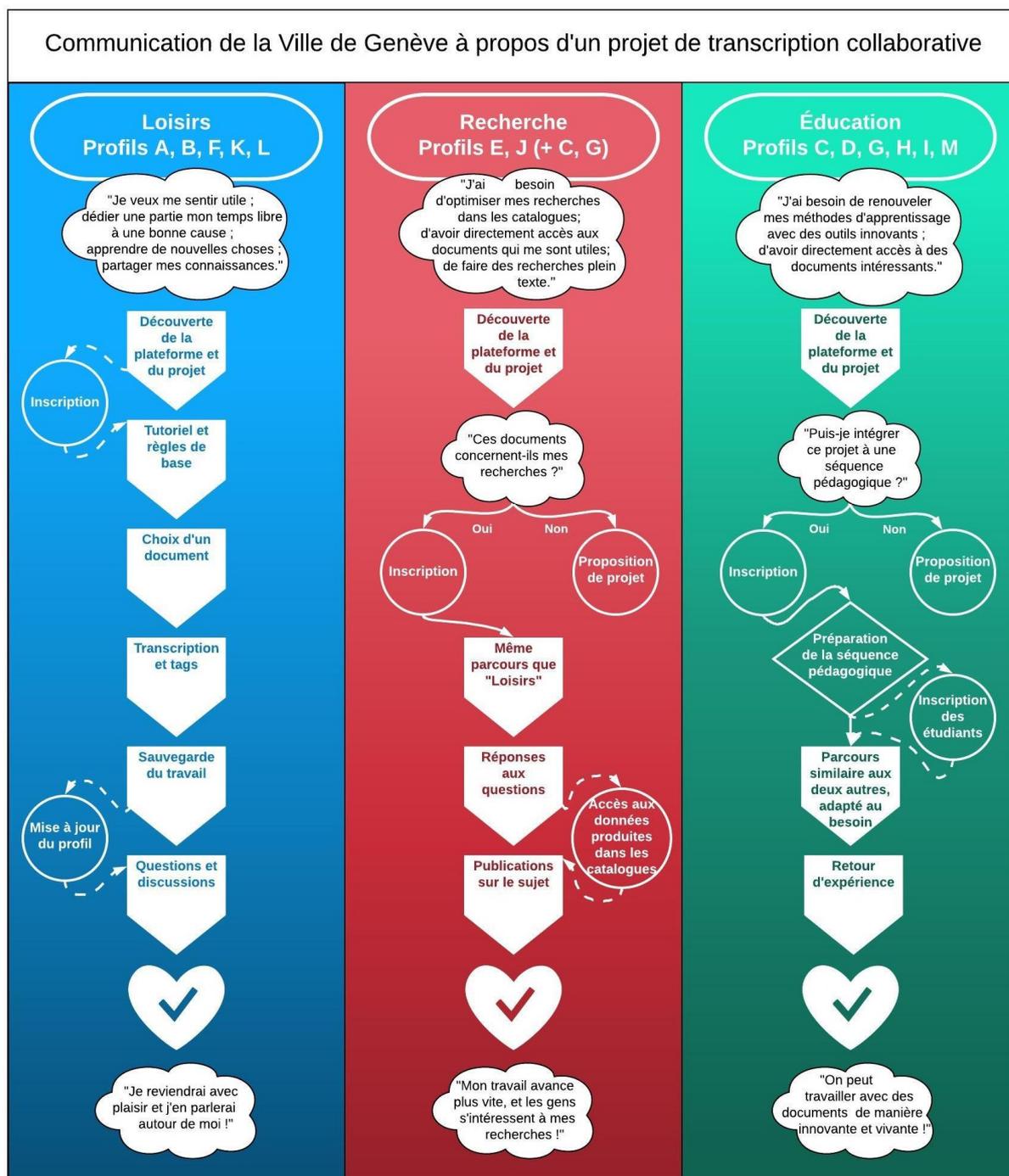
Toutes les informations collectées ont permis d'établir des parcours utilisateurs. Cet outil permet d'envisager les chemins suivis par les futurs usagers sur la plateforme.

Dans le cas présent, trois parcours différents, « Loisirs », « Recherche » et « Éducation », ont été déterminés. À chacun de ces parcours correspond une partie des *persona*, distinguées par leur couleur (bleu, rouge, vert).

Le parcours « Loisirs » représente le parcours d'une personne pour qui découvrir et utiliser la plateforme de transcription collaborative est une fin en soi. À l'inverse, pour les parcours « Recherche » et « Éducation », la plateforme est un moyen permettant d'atteindre un autre but : avancer dans ses recherches (en professionnel ou en amateur), ou enrichir un moment éducatif (en tant qu'enseignant ou apprenant). La *persona* « Bibliothécaire de prison » suit donc le parcours « Éducation », car la plateforme interviendrait dans le cadre de la réinsertion et de la stimulation intellectuelle des détenus, si les conditions d'utilisation le permettent.

Tous les parcours commencent par une communication de la Ville de Genève, car il s'agit d'un nouveau service que les futurs usagers ont peu de chance de découvrir par hasard. Le parcours « Loisirs » représente le parcours classique d'un usager de la plateforme. Les deux autres parcours débutent d'une manière qui leur est propre, mais ils deviennent ensuite similaires au parcours « Loisirs ». À partir de ce point, ils suivent les étapes du parcours « Loisirs », de « Tutoriel et règles de base » à « Questions et discussions », de manière plus ou moins linéaire selon les besoins, avant de continuer et de terminer leur propre chemin.

Figure 28 : Parcours utilisateurs



Les parcours ne sont pas des huis-clos. Un chercheur qui ne trouverait pas de documents utiles à son travail sur la plateforme pourrait tout à fait y revenir dans son temps libre, et suivre alors le parcours « Loisirs ». De même, un enseignant universitaire est généralement un chercheur également, et les parcours rouge et vert sont tous deux envisageables.

Ces parcours utilisateurs permettent néanmoins de mieux envisager les futurs prototypes, les étapes importantes à inclure, et les moments où un usager pourrait abandonner la plateforme si elle ne répond pas à ces besoins. L'option « proposer un projet » est là pour pallier ce risque, mais cela reste un élément dépendant de la stratégie et du niveau de participation de ce projet. Ces parcours devront être mis à jour et détaillés par la suite, une fois le projet lancé, en étudiant le parcours d'utilisateurs réels et les étapes où l'expérience doit être améliorée.

4.8 Prototype et outils

Les étapes précédentes ont mené à la réalisation d'un prototype. Par manque de temps, il s'agit d'une maquette permettant de visualiser l'agencement des divers éléments, et non d'un outil informatique évaluable en conditions réelles. Des outils envisageables pour un futur prototype, ainsi que des lignes directrices de sélection sont néanmoins proposés ici.

4.8.1 Maquette

Pour réaliser la maquette, des schémas ont d'abord été réalisés [sur papier](#), avant de la créer avec l'outil [Moqups](#). L'objectif est de suivre autant que possibles les tendances que ce travail a fait émerger, tout en se tenant à la ligne graphique de Mirabilia. Les limites de Moqups en version gratuite impliquent qu'il n'a pas été possible de créer des pop-ups (pour le tutoriel par exemple), et que tous les éléments n'ont pas pu être ajoutés lors de la première itération.

4.8.1.1 Présentation de la maquette

La maquette comprend sept pages, liées les unes aux autres. Elle commence par une **page d'accueil**, qui reprend la ligne graphique de Mirabilia en annonçant un nouveau projet. L'idée est de proposer un thème parlant et suffisamment vaste pour que chacune des institutions puisse y rattacher un projet de *crowdsourcing* qui concerne ses propres collections.

Figure 29 : Page d'accueil de la maquette



[Lien de l'image](#)

L'autrice et voyageuse genevoise Ella Maillart est une suggestion envisageable *a priori* car, outre la transcription de documents conservés à la BGE, voire aux AVG, des projets de *crowdsourcing* autour de collections de faune (MHN) et de flore (CJB) ou encore d'objets d'art (Ariana, FMAC, MAH, MEG) liés à ses destinations de voyage, pourraient être conçus, mais cela sort du champ de ce travail. En outre, les BM pourraient suggérer des ouvrages sur ce thème, en liant la plateforme à leur blog ou à leur catalogue.

L'icône en haut à droite offre la possibilité de se créer un compte, mais il s'agit d'une décision que devront prendre les responsables du projet. La page d'inscription n'a pas pu être créée.

Les pages suivantes (disponibles en [annexe](#)), sont le **menu**, auquel ont été ajoutées une image dans le thème et la mention « Vos contributions », ainsi que la page de **présentation du projet**. Cette dernière devrait proposer à la fois le contexte et les objectifs du projet, des informations générales sur ce qui est attendu des contributeurs et à quoi serviront ces contributions, et enfin des statistiques permettant d'estimer l'état d'avancement du projet.

Sur cette page quatre onglets sont disponibles, dont un seul est actif, « Transcrire », les autres concernant d'autres types de *crowdsourcing*. Cet onglet mène à la **liste de documents** disponibles pour la transcription, dont l'état d'avancement est indiqué. Des images et métadonnées permettent d'aider au choix. Des filtres de recherche devraient être ajoutés.

Seul le premier document est actif, et il mène à la page **Carnet 1**, qui représente l'outil de transcription. Le **tutoriel et les règles de transcription** sont atteignables par un bouton, qui mène à la page dédiée, mais un pop-up aurait été préférable. Cette page présente un tutoriel vidéo, des règles de transcription de base ainsi que des cas particuliers, qui s'inspire du *Field Guide* de Zooniverse. On peut ensuite revenir à la page précédente.

Le document qui, pour des raisons de disponibilités en ligne, est un écrit de Nicolas Bouvier et non d'Ella Maillart, se trouve à gauche de la page, avec des possibilités de zoomer, retourner l'image etc. La zone de transcription et les zones de tags par catégorie suivent le modèle d'Europeana. Les possibilités de formatage sont simples, avec un bouton pour signaler un problème, et l'on peut sauvegarder sa contribution, ce qui ramène vers le **Profil**.

Cette dernière page, atteignable également depuis l'icône en haut à droite sur les autres pages, permet de voir son profil utilisateur, avec les dernières contributions enregistrées et des statistiques propres (nombre de contributions etc.).

4.8.1.2 Tests

La maquette a été envoyée aux usagers potentiels ainsi qu'aux responsables de Mirabilia pour être testée. Les retours ont été positifs, et les répondants ont apprécié les éléments suivants :

- Visuel sobre et esthétique, dans la ligne de Mirabilia, avec une navigation intuitive
- Outil *a priori* simple d'utilisation, avec diverses possibilités d'enrichissements
- Possibilité de manipuler l'image (zoom etc.)
- Présentation et organisation du tutoriel et des règles de transcription
- Statistiques visibles qui pourront donner envie d'avancer
- Choix du sujet (Ella Maillart, mais aussi Nicolas Bouvier)

Ils notent cependant quelques points d'amélioration :

- Manque de détails quant au contexte du projet et aux consignes
- Manque de liens externes (vers la biographie d'Ella Maillart, par exemple)
- Choix des images et documents ne représentant pas assez de variété d'individus
- Formulations parfois peu explicites
- Incompréhension quant aux onglets « Tagger », « Classifier » et « Lire »
- Position du tutoriel, qui devrait être disponible plus tôt
- Nature des statistiques peu claire
- Incompréhension quant à l'icône de bulle avec un point d'exclamation
- Zones cliquables pas toujours claires
- Flèches de retour qui ne fonctionnent pas très bien
- Boutons « Sauvegarder » et « Soumettre à la relecture » qui ramènent au Profil

Certains de ces problèmes sont directement liés aux limites de l'outil Moqups en version gratuite, par exemple les flèches de retour, et le manque de détails et d'explications, par manque de place. D'autres découlent d'éléments de stratégie du projet que doivent encore déterminer les responsables, d'où un certain flou ou des décisions *a priori* arbitraire. C'est le cas du choix des images, de la nature des statistiques et des boutons « Sauvegarder » et « Soumettre à la relecture » (qui dépendront du système de validation choisi).

D'autres dépendent en revanche d'un manque de clarté au moment de l'envoi de la maquette, et de choix qui méritent d'être reconsidérés. La maquette a donc été mise à jour en ce sens.

4.8.1.3 Itération

Sur la base des retours ci-dessus, des modifications ont été apportées à la maquette. La seconde version, disponible en [annexe](#), est testable en suivant ce [lien](#).

- Page d'accueil
 - Ajout d'un bouton « En savoir plus sur Ella Maillart », qui mène à sa page Wikipédia. Une autre source d'informations est bien sûr envisageable.
 - Le bouton « Plus » a été ôté, du fait de la limite du nombre d'éléments
- Menu
 - Ajout d'un onglet « Proposer un projet », pour faire suite aux parcours utilisateurs. Aucune page n'a pu être créée pour cela, malheureusement.
- Présentation du projet
 - La formulation « créez-vous un compte » a été changée en « créez votre propre compte », sur les conseils d'un usager.

- Transcrire
 - Un bouton « Comment s’y prendre ? » (formulation suggérée par un usager) a été ajouté, pour amener directement au tutoriel Un tutoriel pop-up demeure cependant un choix plus judicieux, *a priori*.
 - Le titre Carnet 1 est cliquable, alors que seul l’image l’était auparavant
 - Le bouton « Plus » a été ôté, du fait de la limite du nombre d’éléments
- Carnet 1
 - La formulation concernant les tags a été modifiée, sur conseil d’un membre du Comité éditorial
 - Une courte explication a été ajoutée à côté de l’icône de bulle avec un point d’exclamation
- Tutoriel et règles de transcription
 - Les flèches de retour sont liées à deux pages différentes, et non plus à la page d’accueil. Ces flèches demeurent un problème dans la maquette, du fait des limites de Moqups.
- Profil
 - Le logo de Mirabilia a été ôté, du fait de la limite du nombre d’éléments

D’autres itérations et tests sont bien entendu nécessaires, mais il est préférable de réaliser d’abord un prototype d’outil de transcription afin de pouvoir lier les deux. La maquette donne cependant des repères intéressants quant aux attentes des futurs usagers et aux éléments minimaux requis pour la plateforme.

Sur cette base et en fonction des résultats des étapes précédentes, on peut définir les éléments suivants pour l’élaboration de la future plateforme :

- Plateforme minimale
 - Contexte détaillé et consignes claires
 - Libre choix des documents
 - Indication de l’état d’avancement des transcriptions
 - Accès aux documents déjà transcrits
 - Tutoriel et règles de transcription à portée de main
 - Possibilité de se créer un compte pour retrouver ses transcriptions
 - Possibilité de signaler si l’on a un doute sur un mot ou un caractère
 - Système de formatage du texte (quel qu’il soit)
 - Système de validation explicite (quel qu’il soit)

- Les « plus »
 - Tags divers
 - Choix de document aléatoire possible
 - Contribution en anonyme possible (si la modération le permet)
 - Espaces de discussion, de questions
- Le « top »
 - Pouvoir proposer un projet, un corpus de documents

Un système de formatage du texte plus complexe que celui de la maquette ainsi qu'un processus de validation différent peuvent être envisagés, selon l'outil choisi.

4.8.2 Outils envisageables

La maquette ci-dessus permet d'avoir une idée du visuel envisageable et des fonctionnalités minimales nécessaires, cependant le temps a manqué pour réaliser et faire tester des prototypes fonctionnels. Il faut pour cela commencer par sélectionner un outil existant.

Un article de blog de Ben Brumfield, daté de 2012, renvoie à un document Google Sheet comprenant une liste d'outils de transcription collaborative et leur descriptif selon certains critères (Brumfield 2012b). Ce document, disponible en [annexe](#), a été mis à jour plusieurs fois par l'auteur ou d'autres contributeurs, et comporte aujourd'hui quarante-cinq outils différents. Cependant, les derniers ajouts datent, et certains outils récents, comme *TACT* ou *Concordia*, n'y figurent pas. Le choix est donc vaste pour qui veut monter un tel projet, d'autant que, comme le signale Ben Brumfield il n'existe pas d'outil « one size fits all » (Brumfield 2013), et la sélection dépend des spécificités techniques et stratégiques du projet.

Dans un premier temps, des conditions *sine qua non* ont été déterminées, sur la base des étapes précédentes et de différents échanges avec les responsables de Mirabilia :

- L'interface doit être en français
- Le projet ne doit pas être hébergé sur une plateforme externe gérée par un tiers, et il faut donc privilégier des outils *open source*
- Les transcriptions produites doivent pouvoir être exportées sur les catalogues pour permettre de les exploiter
- On doit pouvoir proposer des fonds variés et multiples
- L'utilisateur doit pouvoir choisir le document qu'il souhaite transcrire, mais une option « choix aléatoire » peut être envisagée

La demande d'accès au document Google Sheet a été l'occasion d'une rencontre en ligne avec Ben et Sara Brumfield, lors de laquelle ils ont relevé les deux facteurs suivants :

- Le type de documents traités
- Les modalités d'encodage du texte

Pour le type de documents, la distinction repose essentiellement sur la notion de documents structurés ou non structurés, évoquée [plus haut](#). Les outils sont souvent mieux adaptés à l'un ou à l'autre, à l'instar de *Scribe*, une référence pour les documents structurés (Brumfield 2020), mais tous permettent le traitement des deux types, dans une certaine mesure.

Quant aux modalités d'encodage de texte, certains outils proposent une transcription avec peu ou pas de formatage, comme *Scripto*, et d'autres un balisage XML plus complexe, comme *Transcription Desk*. D'après Ben et Sara Brumfield, le choix dépend d'une part du public et des objectifs visés, car le balisage XML peut être complexe à maîtriser mais permet des transcriptions bien plus fines, et d'autre part des catalogues et bases de données dans lesquels seront exportées les transcriptions. Si les outils de recherche ne supportent pas le XML, rien ne sert de faire travailler les contributeurs avec ce langage.

Concernant ce second point, une partie des institutions genevoises est en train de changer de système de catalogue collectif, avec le projet [SLSP](#). L'équipe en charge a été contactée pour plus d'informations, mais une discussion n'a pu encore avoir lieu. Ces institutions ont en outre des catalogues et bases de données spécialisées qui leur sont propres et qui n'ont pas tous les mêmes fonctionnalités. Il faudra donc sélectionner un outil qui sied au plus grand nombre.

Sur la base du document Google Sheet, de la littérature et des divers tests effectués, six outils ont été retenus, dont quatre sont très présents dans la littérature et deux sont récents. Tous remplissent *a priori* les conditions *sine qua non*, mais le temps a manqué pour une étude plus approfondie, qui sera nécessaire à terme. Il faudra en outre s'assurer que l'outil soit compatible avec le CMS de Mirabilia, Drupal. Ces outils sont présentés dans le tableau ci-dessous :

Tableau 2 : Comparatif d'outils de transcription collaborative

	Adapté surtout à des documents...	Type d'encodage	Autres éléments
FromThePage	Non-structurés	Diverses possibilités	Tags possibles
Scribe	Structurés	Formatage simple, avec export en CSV	Vérification <i>multi-track</i>
Scripto	Non-structurés	Formatage simple	Fonctionnait avec Drupal, mais les informations n'ont pas été mise à jour récemment
Transcription Desk	Non-structurés	XML	Basé sur MediaWiki
Concordia	Non-structurés	Formatage simple	Maintenu par la LoC
TACT	Non-structurés	XML	Projets multiples possibles

Ce tableau se base sur les informations qui ont pu être récoltées à leur sujet, mais il sera nécessaire d'aller plus loin en regardant notamment si le système de validation est prédéfini, comme c'est le cas de *Scribe*, et s'il est possible de tagger les documents, comme c'est le cas de *FromThePage*. En définitive, le choix dépendra de la stratégie qui sous-tendra le projet, mais aussi, et c'est essentiel, de l'avis et des compétences de l'équipe technique du DCTN et du fournisseur web de Mirabilia.

Dans tous les cas, il serait judicieux, voire impératif, d'éprouver le ou les outils adéquats afin de créer rapidement des prototypes minimaux pouvant être testés par les usagers potentiels et par les tenants du projet. Ceci permettra de déterminer quelle direction suivre, et d'améliorer le prototype par itération. Ce processus devra continuer même après le lancement du projet, et ses responsables devront être prêts à continuer l'amélioration de l'outil par itération au fur et à mesure des retours des usagers et de l'émergence de nouveaux besoins.

4.8.3 Informations récentes et proposition de solutions⁷

À la suite du rendu de ce Travail de Master, et en prévision de la défense à venir, des rencontres ont eu lieu avec diverses personnes contactées préalablement, mais qui n'étaient jusque-là pas disponibles. Ces personnes, toutes employées par des institutions culturelles genevoises, ont relevé une spécificité de Mirabilia qui n'était pas évidente à repérer : les notices d'œuvres qui s'y trouvent sont ajoutées manuellement par les collaborateurs, et ne sont donc pas directement liées aux catalogues et bases de données documentaires.

Ce choix a plusieurs implications stratégiques et techniques, et ce notamment dans le cas de la transcription collaborative. En effet, ce type de projets nécessite non seulement un très grand nombre de notices, dont l'ajout manuel serait trop chronophage, mais aussi une mise à jour régulière des notices une fois les transcriptions validées. Un lien direct au catalogue simplifierait ce processus.

Deux solutions se présentent alors. La première serait de suivre le modèle de plateformes telles que Zooniverse, qui ne sont pas directement liées à des catalogues, mais sur lesquelles il est possible d'importer un grand nombre de documents, pour ensuite réexporter les métadonnées créées, au format CSV par exemple. Il faut cependant que l'architecture de Mirabilia permette cela.

La seconde solution serait de faire de Mirabilia non pas une vitrine, mais une porte d'entrée. La plateforme pourrait ainsi conserver son rôle de communication, de valorisation et d'éditorialisation, tout en étant le premier point d'entrée pour les usagers vers les collections et les projets de *crowdsourcing*. Ces derniers seraient développés non pas sur la plateforme elle-même, mais en lien avec les catalogues et bases de données *ad hoc*. Mirabilia pourrait alors non seulement en faire la promotion, mais aussi s'occuper de la coordination des projets entre les institutions, voire de la mutualisation de leurs travaux, par exemple en les encourageant à collaborer au développement d'outils de *crowdsourcing*.

La figure ci-dessous est une version modifiée de la maquette réalisée au préalable, afin de représenter cette seconde solution de manière plus visuelle.

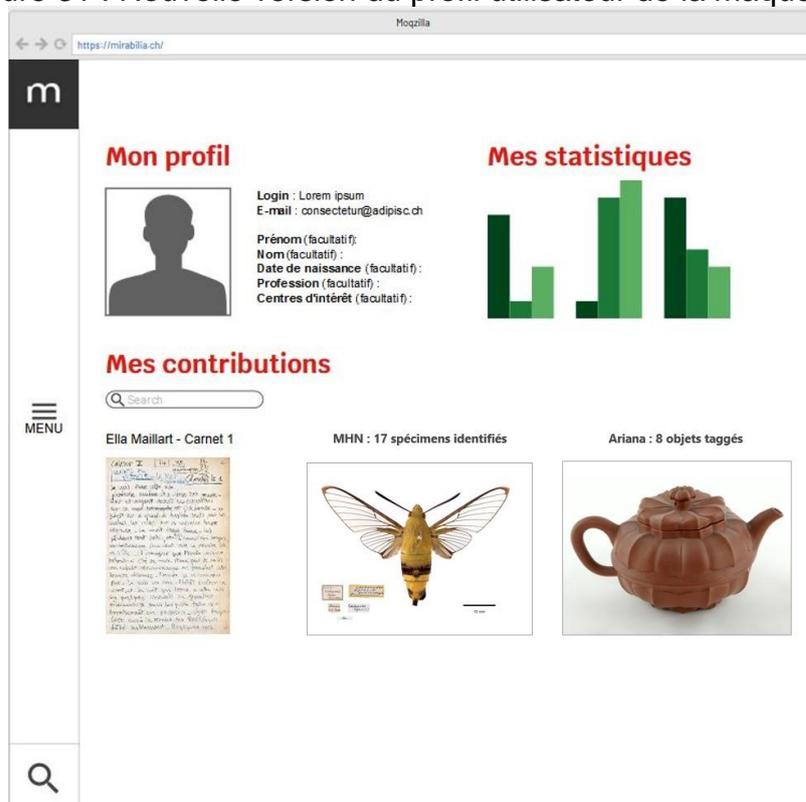
⁷ Cette partie a été ajoutée à la suite de la défense de ce TM, qui a eu lieu le 07.09.2020.

Figure 30 : Nouvelle version de la page d'accueil de la maquette



Il serait en outre intéressant d'avoir un compte utilisateur unique pour l'ensemble des projets, ce qui permettrait à l'utilisateur de retrouver toutes ses contributions en un même lieu. Mirabilia serait ainsi non seulement la porte d'entrée, mais aussi la porte de sortie. La figure ci-dessous présente une version modifiée de la maquette allant dans ce sens.

Figure 31 : Nouvelle version du profil utilisateur de la maquette



4.9 Marche à suivre

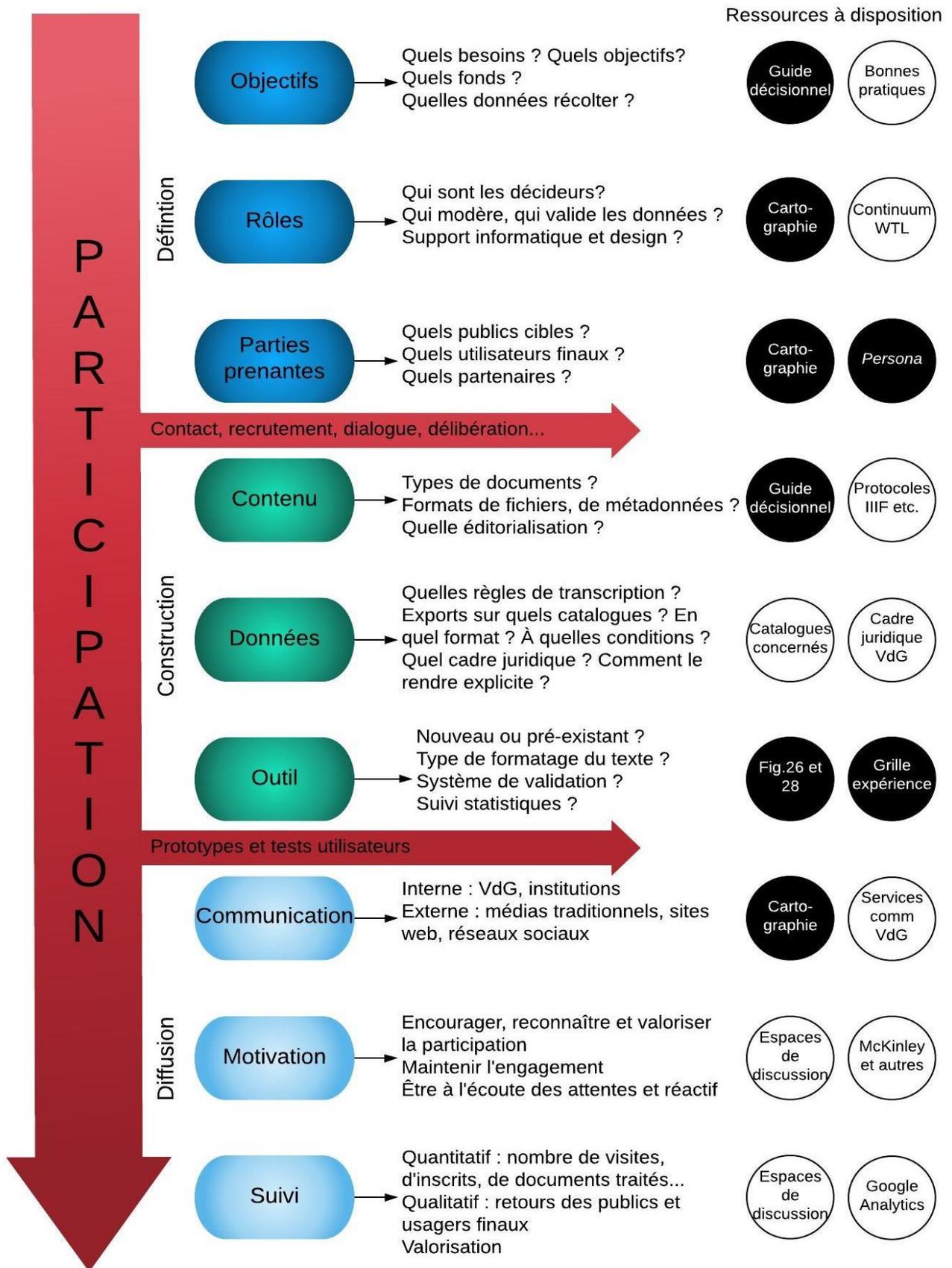
Pour guider la mise en place de futurs projets, une marche à suivre générale et généralisable a été réalisée, sur la base des informations récoltées au cours de ce projet et de la checklist d'Arkhênum (Arkhênum [sans date]).

Cette marche à suivre comporte neuf étapes principales, divisées en trois groupes : définition, construction et diffusion. À chaque étape correspond un certain nombre de questions à se poser ou d'éléments à prendre en compte pour la bonne marche du projet. À droite se trouvent des ressources pouvant servir de base, parmi lesquelles des éléments internes à ce travail (en noir), et des ressources externes (en blanc). Le tout repose sur la participation, qui fait office d'épine dorsale. La participation peut intervenir à tout moment, en amont par exemple, si le projet est porté par un groupe d'utilisateurs, ou plus tard, selon les cas, mais elle doit rester au cœur de la réflexion globale.

Pour des raisons de place, on a utilisé l'abréviation « VdG » pour Ville de Genève, et « Fig. 26 et 28 » pour faire référence à la [figure 26](#) et à la [figure 28](#), aux pages 49 et 54 de ce travail.

Cette marche à suivre est une première ébauche, et elle devra nécessairement être mise à jour au fur et à mesure des expériences, des réalisations, des succès et des échecs.

Figure 32 : Marche à suivre pour de futurs projets



5. Recommandations

Ce travail a fait émerger des lignes directrices et des éléments à prendre en compte, que cette partie résume sous forme de recommandations techniques, qui concernent la réalisation de la plateforme elle-même, et stratégiques, liées au projet dans son ensemble.

5.1 Recommandations techniques

5.1.1 Qualité et interopérabilité

Qu'il s'agisse des documents et métadonnées mis à disposition ou des métadonnées sociales créées par les usagers, la qualité et l'interopérabilité sont essentielles. On peut les assurer en amont en se tenant à certains protocoles internationaux, tels que [IIIF](#) et [OAI-PMH](#).

En ce qui concerne les métadonnées sociales, des règles claires doivent être édictées. Ces règles devraient être suffisamment généralisables pour être appliquées à tout type de manuscrits, et être validées par des experts. Mme Paule Hochuli, conservatrice responsable des manuscrits à la BGE, recommande de retenir au moins deux règles « minimales » : écrire ce que l'on voit, et rétablir la ponctuation et l'accentuation. En plus de ces règles, le formatage des transcriptions par les usagers doit se baser sur un langage suffisamment aisé à prendre en main et compatible avec les catalogues et bases de données dans lesquelles elles seront exportées. Enfin, un processus de vérification et de validation efficace et adapté aux objectifs du projet doit impérativement être mis en place.

5.1.2 Cadre juridique

Pour éviter tout écueil, le cadre juridique qui entoure le projet, la participation des usagers et les données qu'ils produisent doit être clairement défini et explicité. Outre les CGU, il serait judicieux de préciser les éléments importants, par exemple à l'inscription, avec une case à cocher. Les données produites devraient tomber dans le domaine public ou être sous une licence [Creative Commons](#) adéquate. Enfin, les documents à disposition doivent impérativement pouvoir être lus et transcrits, voire téléchargés. Pour des documents récents, un accord avec les ayants droit peut être envisagé.

5.1.3 Communication et collaboration

S'agissant d'un projet collaboratif, la communication entre les usagers devrait être rendue possible, voire encouragée, grâce à des canaux de communication internes à la plateforme. Ceux-ci pourraient également permettre des échanges avec des experts et avec les responsables du projet. Une modération active mais discrète serait de mise. Il vaudrait mieux éviter de pousser les usagers à la concurrence, pour conserver un environnement sain, mais des compétitions amicales, comme des « transcribathons », peuvent être envisagées.

5.1.4 « Construire pour » et « construire avec »

Pour correspondre au mieux à son public, la plateforme doit s'accorder avec eux autant que possible. De ce fait, il est essentiel d'intégrer les usagers aux discussions, délibérations et tests de prototypes en amont du projet, mais aussi d'être à l'écoute de leurs retours pendant et après son déroulement. Il faut être prêt à faire machine arrière ou à modifier l'un ou l'autre élément afin de s'adapter aux usagers, qui savent généralement mieux que personne ce qui leur convient.

5.2 Recommandations stratégiques

5.2.1 Le parapluie et les baleines

Monter un projet commun à toutes les institutions culturelles de la Ville de Genève est un réel défi, tant leurs collections et publics sont variés. La transcription collaborative de manuscrits ne suffit pas à couvrir le champ des possibles qu'offre cette diversité. Il faut donc penser les projets de *crowdsourcing* sur Mirabilia comme un parapluie, avec un grand thème auquel se rattachent des projets indépendants, les baleines, chacun lié à une institution et ses collections, avec ses objectifs et publics propres. Le thème des voyages d'Ella Maillart, proposé [plus haut](#), en est un exemple. Dans ce cadre, la transcription collaborative serait une des solutions de *crowdsourcing*, au sein d'un écosystème plus vaste, mais non la seule.

5.2.2 La vitrine et la porte d'entrée⁸

Comme indiqué précédemment, il serait préférable que les projets de *crowdsourcing* soient construits en lien direct avec les catalogues et bases de données documentaires, et non pas sur Mirabilia, avec des notices importées manuellement. Avec cette solution, Mirabilia tiendrait alors le rôle d'agrégateur de notices, de projets de *crowdsourcing*, mais aussi, pourquoi pas, d'autres éléments, tels que les billets de blog des différentes institutions, que l'on pourrait alors retrouver en un seul lieu. La plateforme deviendrait ainsi le premier point d'entrée pour les usagers, mais aussi, idéalement, leur porte de sortie, si l'on parvient à créer des comptes utilisateurs uniques, sur lesquels seront rassemblées les diverses contributions de chacun.

5.2.3 Une communauté plutôt qu'une foule

Dans le domaine culturel, on parle plus volontiers de *community sourcing* que de *crowdsourcing*, et en effet, mieux vaut cibler et créer des communautés d'intérêt, motivées et soudées, plutôt que d'espérer attirer des foules innombrables. Rappelons-le, le projet de correction d'OCR e-newspaperarchives.ch de la BN avait attiré un peu plus de cent vingt personnes, et il est peu probable qu'un projet local ait un rayonnement beaucoup plus grand. Mieux vaut une communauté enthousiaste et active qu'une foule qui se désintéresse vite.

5.2.4 Ouvrir la porte...

Choisir le participatif, c'est ouvrir la porte à de nouveaux usages, de nouvelles relations avec son public, et de nouveaux rôles pour chacun. Rappelons la phrase de Graham Black :

“Thus, developing a participatory museum is not focused on increasing visitor numbers but, rather, on building lasting relationships between a museum and its audiences and/or communities and converting audiences into cultural participants, thereby increasing the museum’s relevance to the communities it serves” (Black 2018, p. 317)

Les bénéfices sont nombreux, mais ce choix implique également d'accepter les changements qui l'accompagnent. Les professionnels doivent être prêts à laisser une place aux amateurs, à leur faire confiance et à reconnaître la légitimité de leurs savoirs et la valeur de leurs contributions. Une communication et une médiation bien pensées permettront dans tous les cas de préserver l'image des institutions comme garantes d'une information de qualité.

⁸ Cette partie a été ajoutée à la suite de la défense de ce TM, qui a eu lieu le 07.09.2020.

5.2.5 ...et sortir de la pièce

Un projet participatif suppose d'intégrer les usagers, de leur donner la place de s'exprimer et d'agir, d'être à leur écoute et ouverts à leurs demandes. Ces éléments peuvent se développer avec le temps et prendre une importance croissante, si on leur en laisse la possibilité.

Au fur et à mesure, la participation peut ainsi évoluer en suivant, par exemple, le [Continuum](#) du projet WTL, et en s'inspirant de la phrase de Peter Samis : « We opened the door to let visitors in... then we left the room » (cité dans Ridge 2013, p. 3). Des projets de *crowdsourcing* pourront alors être portés et gérés par des groupes d'usagers, et la plateforme pourra être modelée selon leurs besoins et désirs.

Dans un tel contexte, les institutions instigatrices du projet Mirabilia seront amenées à agir en coulisses, en rendant disponibles leurs collections, en mettant en lien des personnes aux parcours variés, en leur permettant de se sentir légitimes dans leurs contributions et en les invitant à s'appropriier la plateforme et le patrimoine genevois.

6. Conclusion

Cette étude a permis de poser les bases d'un projet de transcription collaborative de manuscrits en Ville de Genève, sur la plateforme Mirabilia. En questionnant la demande institutionnelle d'une part, et en interrogeant des usagers potentiels d'autre part, elle a fait émerger des lignes directrices et des recommandations qui pourront faire office de cadre.

Ce travail s'est appuyé sur des méthodes tirées du *Design Thinking* et du UX. Il s'est construit en plusieurs phases. La première a été une recherche et des observations personnelles via des lectures, des tests de plateformes existantes et des rencontres avec des experts. La suivante a consisté à mieux comprendre les besoins et attentes des tenants du projet *via* des questionnaires, à déterminer des publics cibles, et à créer des *persona* pour pouvoir recruter des usagers potentiels. La troisième phase visait à recueillir les opinions de ces usagers potentiels sur la base de tests de plateformes et de questionnaires, et à les impliquer dans la définition du projet *via* un tri de cartes. Enfin, la dernière phase a consisté à définir des parcours utilisateurs, à créer un premier prototype, à le faire tester, et à proposer des lignes directrices pour la suite, sous forme de suggestions d'outils et d'une marche à suivre.

Ces étapes ont démontré l'importance de confronter la vision des tenants du projet aux besoins et attentes du public, car ils ne coïncident pas toujours. En consultant les usagers potentiels, le projet peut prendre des directions différentes et un sens nouveau. On en retient tout d'abord l'importance, pour les amateurs, de la notion de communauté, de pouvoir se rencontrer en groupes d'intérêt et de collaborer à un projet de grande envergure, sur des thématiques très variées. Il est cependant essentiel pour ces personnes de se sentir légitimes dans leur participation et de voir que leurs contributions sont utiles. Concernant les enseignants et les chercheurs, il est impératif d'aller chercher le dialogue avec eux. Leur intérêt pour la plateforme dépend du lien avec leurs sujets de travail, et s'ils ne sont pas consultés, s'ils n'ont pas leur mot à dire sur le choix de thématiques et de fonds, ils risquent de désertir la plateforme.

Il faut donc que les tenants du projet soient prêts à intégrer les usagers dans la construction de ce projet participatif et à leur laisser de la place, sans pour autant craindre pour l'image de leur institution. Il faut pour cela prendre le temps de définir des objectifs précis pour éviter l'écueil d'un projet « prétexte », sans autre but que de suivre la tendance à la participation. Il faudra ensuite continuer le processus entamé en créant des prototypes d'outils et en les faisant tester à des usagers potentiels, afin de s'assurer qu'ils répondent à leurs besoins et attentes. Le lancement du projet devra s'appuyer sur une communication à grande échelle, impliquant les médias traditionnels, mais ciblant en particulier des communautés spécifiques.

Les résultats de ce travail découlent d'une méthodologie qui a dû être adaptée aux conséquences de la crise sanitaire de 2020. De ce fait, certaines étapes importantes n'ont pas pu être réalisées dans les temps, et certains biais persistent. Si les principes du *Design Thinking* avaient pu être appliqués autrement qu'à distance, des éléments différents, plus spécifiques encore au contexte genevois, seraient peut-être ressortis. Cette méthode fait en effet ses preuves depuis des années et demeure un choix idéal pour la construction d'un projet participatif. Il serait donc judicieux, par la suite, si la situation le permet, d'appliquer cette méthode en présentiel pour en exploiter tout le potentiel. Il faudrait en outre être prêt à faire évoluer les outils produits dans ce travail et à réajuster la direction proposée, au fil des expériences et des réalisations.

Bibliographie

ALAM, Sultana Lubna, SUN, Ruonan et CAMPBELL, John, 2020. Helping Yourself or Others ? Motivation Dynamics for High-Performing Volunteers in GLAM Crowdsourcing. *Australasian Journal of Information Systems* [en ligne]. 18 mai 2020. Vol. 24. [Consulté le 17 juin 2020]. Disponible à l'adresse :

<https://journal.acs.org.au/index.php/ajis/article/view/2599>

ALCALÁ PONCE DE LEÓN, Mireia, 2015. Crowdsourcing in the memory institutions: mass transcriptions. *BiD: textos universitaris de biblioteconomia i documentació* [en ligne]. Décembre. N° 35. [Consulté le 17 juin 2020]. Disponible à l'adresse : <http://bid.ub.edu/en/35/alcala.htm>

ALDANA, Mélissa, DROMER, Vincent et LEMÉNI, Yoann, 2019. *Passez au design thinking : penser, construire et mener vos premiers ateliers de cocréation*. Paris : Editions Eyrolles. ISBN 978-2-212-57059-5.

ARKHÊNUM, [sans date]. *La checklist pour un projet de valorisation réussi*. Bordeaux : Citron Pressé. Document interne.

BATS, Raphaëlle (éd.), 2015. Construire des pratiques participatives dans les bibliothèques. Villeurbanne : Presses de l'enssib. La Boîte à outils. ISBN 978-2-37546-069-6.

BATS, Raphaëlle, 2019. De la participation à la mobilisation collective, la bibliothèque à la recherche de sa vocation démocratique [en ligne]. Paris : Université Paris Diderot. Thèse. [Consulté le 23 juin 2020]. Disponible à l'adresse : <https://hal.archives-ouvertes.fr/tel-02465951>

BAUMANN, Jan, 2018. Projet de transcription collaborative sur e-manuscripta.ch. *Infoclio.ch* [en ligne]. 18 septembre 2018. [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://www.infoclio.ch/fr/projet-de-transcription-collaborative-sur-e-manuscriptach>

BEUDON, Nicolas, 2015a. Qu'est-ce que le design thinking ? *Le recueil factice* [en ligne]. 18 mars 2015. [Consulté le 23 avril 2020]. Disponible à l'adresse : <http://nicolas-beudon.com/2015/03/18/design-thinking-1/>

BEUDON, Nicolas, 2015b. Comment introduire le design thinking en bibliothèque ? *Le recueil factice* [en ligne]. 20 mars 2015. [Consulté le 23 avril 2020]. Disponible à l'adresse : <http://nicolas-beudon.com/2015/03/20/design-thinking-2/>

BEUDON, Nicolas, 2017a. Le vocabulaire du design thinking. *Le recueil factice* [en ligne]. 8 septembre 2017. [Consulté le 23 avril 2020]. Disponible à l'adresse : <http://nicolas-beudon.com/2017/09/08/vocabulaire/>

BEUDON, Nicolas, 2017b. Les 8 commandements de l'UX en bibliothèque. *Le recueil factice* [en ligne]. 14 septembre 2017. [Consulté le 23 avril 2020]. Disponible à l'adresse : <http://nicolas-beudon.com/2017/09/14/ux/>

BEUDON, Nicolas, 2019a. *Design Thinking et design de service* [document PDF]. Support de cours. Cours « Gestion de services d'information », Haute école de gestion de Genève, Master IS, semestre d'automne 2019.

BEUDON, Nicolas, 2019b. Travailler autrement avec la pensée visuelle et les canevas. *Le recueil factice* [en ligne]. 22 septembre 2019. [Consulté le 23 avril 2020]. Disponible à l'adresse : <http://nicolas-beudon.com/2019/09/22/travailler-autrement-avec-la-pensee-visuelle-et-les-canevas/>

- BISO, Stéphane et LE NAOUR, Marjorie, 2017. *Design thinking : accélérez vos projets par l'innovation collaborative*. Malakoff: Dunod. ISBN 978-2-10-076729-8.
- BLACK, Graham, 2018. Meeting the audience challenge in the 'Age of Participation'. *Museum Management and Curatorship* [en ligne]. 4 juillet 2018. Vol. 33, n° 4, pp. 302-319. [Consulté le 25 juin 2020]. Disponible à l'adresse : <https://www.tandfonline.com/doi/full/10.1080/09647775.2018.1469097>
- BLASER, Lucinda, 2014. Old Weather: Approaching Collections from a Different Angle. In : RIDGE, Mia, 2014. *Crowdsourcing our Cultural Heritage*. Farnham. : Ashgate, pp. 45-55. ISBN 978-1-4724-1022-1.
- BODER, Alexandre, DUBOSSON, Françoise, ANDERFUHREN, Sandrine et FUSS, Megan, 2016. *Vu, connu, reconnu, et commenté, réutilisé, détourné, enrichi : les pratiques de valorisation des collections numériques dans les institutions culturelles : état des lieux et recommandation*. Genève : Haute école de gestion, décembre 2016. Document interne
- BODER, Alexandre et DUBOSSON, Françoise, 2020. *Ça m'intéresse... Je participe ! Pratiques participatives et transcription collaborative adaptées à la plateforme Mirabilia : état des lieux et recommandations*. Genève : Haute école de gestion, août 2020. Document interne
- BRABHAM, Daren C., 2013. *Crowdsourcing*. Cambridge, MA ; Londres : The MIT Press. The MIT Press essential knowledge series. ISBN 978-0-262-51847-5.
- BROKFELD, Jens, 2012. *Die digitale Edition der „preußischen Zeitungsberichte“: Evaluation von Editionsworkzeugen zur nutzergenerierten Transkription handschriftlicher Quellen* [en ligne]. Potsdam : Fachhochschule Potsdam. Travail de Master. [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://opus4.kobv.de/opus4-fhpotsdam/frontdoor/index/index/docId/331>
- BRUMFIELD, Ben W., 2011. 2010 : The Year of Crowdsourcing Transcription. *FromThePage Blog* [en ligne]. 2 février 2011. [Consulté le 16 juillet 2020]. Disponible à l'adresse : <https://content.fromthepage.com/2011/02/>
- BRUMFIELD, Ben W., 2012a. Quality Control for Crowdsourced Transcription. *FromThePage Blog* [en ligne]. 5 mars 2012. [Consulté le 22 juin 2020]. Disponible à l'adresse : <https://content.fromthepage.com/2012/03/>
- BRUMFIELD, Ben W., 2012b. Crowdsourced Transcription Tool List. *FromThePage Blog* [en ligne]. 11 avril 2012. [Consulté le 22 juin 2020]. Disponible à l'adresse : <https://content.fromthepage.com/2012/04/>
- BRUMFIELD, Ben W., 2013. Choosing Crowdsourced Transcription Platforms at SSA 2013. *FromThePage Blog* [en ligne]. 24 mai 2013. [Consulté le 24 juin 2020]. Disponible à l'adresse : <https://content.fromthepage.com/2013/05/>
- BRUMFIELD, Ben W., 2020. The Decade in Crowdsourcing Transcription. *Crowdsourcing Week* [en ligne]. 13 janvier 2020. [Consulté le 24 juin 2020]. Disponible à l'adresse : <https://crowdsourcingweek.com/blog/decade-in-crowdsourcing-transcription/>
- CAUSER, Tim, GRINT, Kris, SICHANI, Anna-Maria et TERRAS, Melissa, 2018. « Making such bargain » Transcribe Bentham and the quality and cost-effectiveness of crowdsourced transcription. *Digital Scholarship in the Humanities* [en ligne]. 15 janvier 2018. Vol. 33. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://discovery.ucl.ac.uk/id/eprint/10039865/>

CHAUMIER, Serge et FRANÇOISE, Camille, 2014. Museomix : l'invention d'un musée du XXI^e siècle. *La Lettre de l'OCIM* [en ligne]. 1 novembre 2014. N° 156, pp. 7-11. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://ocim.fr/lettre/dispositifs-mediation-innovants-recensement-herbiers/>

CHRON, Otto et SUNDELL, Sami, 2011. Digitalkoot : Making Old Archives Accessible Using Crowdsourcing. In : *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, 8 août 2011* [en ligne]. [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/view/3813/4246>

CHUPIN, Lisa, 2016. Documentarisation participative et médiation du patrimoine scientifique numérisé. Le cas des herbiers. *Études de communication. Langages, information, médiations* [en ligne]. 1 juin 2016. N° 46, p. 33-50. [Consulté le 17 juin 2020]. Disponible à l'adresse : <http://journals.openedition.org/edc/6499>.

DIAS DA SILVA, Patrícia, HEATON, Lorna et MILLERAND, Florence, 2017. Une revue de littérature sur la « science citoyenne » : la production de connaissances naturalistes à l'ère numérique. *Natures Sciences Sociétés* [en ligne]. 2017. Vol. 25, n° 4, pp. 370-380. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://www.cairn.info/revue-natures-sciences-societes-2017-4-page-370.htm>

ESTELLÉS-AROLAS, Enrique et GONZÁLEZ-LADRÓN-DE-GUEVARA, Fernando, 2012. Towards an integrated crowdsourcing definition. *Journal of Information Science* [en ligne]. 1 avril 2012. Vol. 38, n° 2, p. 189-200. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://journals.sagepub.com/doi/10.1177/0165551512437638>

ESTERMANN, Beat, 2013. *Swiss Heritage Institutions in the Internet Era. Results of a pilot survey on open data and crowdsourcing* [en ligne]. Bern: Bern University of Applied Sciences, E-Government Institute. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://arbor.bfh.ch/9503/>

ESTERMANN, Beat, 2015. Open data et crowdsourcing : un état des lieux du point de vue des musées. *La Lettre de l'OCIM. Musées, Patrimoine et Culture scientifiques et techniques* [en ligne]. 1 novembre 2015. N° 162, pp. 41-46. [Consulté le 17 juin 2020]. Disponible à l'adresse : <http://journals.openedition.org/ocim/1597>

EUROPEANA, Bienvenue sur Europeana Collections. *Europeana* [en ligne]. 16 juillet 2020. [Consulté le 16 juillet 2020]. Disponible à l'adresse : <https://www.europeana.eu/fr/about-us>

FERRITER, Meghan, ZWAARD, Kate, KAMLLEY, Elaine, STOREY, Rosie, ADAMS, Chris, ALGEE, Lauren, HYNING, Victoria Van, BRESNER, Jamie, POTTER, Abigail, JAKEWAY, Eileen et BRUNTON, David, 2019. "With One Heart": Agile approaches for developing Concordia and crowdsourcing at the Library of Congress. *The Code4Lib Journal* [en ligne]. 5 novembre 2019. N° 46. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://journal.code4lib.org/articles/14901?loclr=blogsig>

GLASER, David, 2020. Rapport de sondage de notrehistoire.ch. *Issuu* [en ligne]. 19 mars 2020. [Consulté le 13 avril 2020]. Disponible à l'adresse : https://issuu.com/david.glaser/docs/rapport_de_sondage_notrehistoire.ch_2020

GRANVAL, Aurore et D'ANS, Micha, 2019. PLUME : la bibliothèque met en ligne ses collections patrimoniales. *EPFL Magazine* [en ligne]. Mars 2019. N° 24, p. 29. [Consulté le 17 juin 2020]. Disponible à l'adresse : https://epfl.magnum3.ch/patrons/epfl/issues/6/issue_fr.pdf

HOWE, Jeff, 2006a. The Rise of Crowdsourcing. *Wired* [en ligne]. 1^{er} juin 2006. N° 14, p. 5. [Consulté le 23 juin 2020]. Disponible à l'adresse : <https://www.wired.com/2006/06/crowds/>

HOWE, Jeff, 2006b. Crowdsourcing: A Definition. *Crowdsourcing* [en ligne]. 2 juin 2006, 10 :30. [Consulté le 23 juin 2020]. Disponible à l'adresse : https://www.crowdsourcing.com/cs/2006/06/crowdsourcing_a.html

HOLLEY, Rose, 2010. Crowdsourcing: How and why should libraries do it? *D-Lib Magazine* [en ligne]. 15 mars 2010. [Consulté le 17 juin 2020]. Disponible à l'adresse : <http://eprints.rclis.org/14360/>

ICOM, 2019. Définition du musée. *ICOM* [en ligne]. [Consulté le 17 mai 2020] Disponible à l'adresse : <https://icom.museum/fr/ressources/normes-et-lignes-directrices/definition-du-musee>

IDEO, 2016. *Le Design Thinking en bibliothèque : un kit pratique pour la conception de projets centrés sur les usagers* [en ligne]. Bill and Melinda Gates foundation : 2016 [consulté le 21 avril 2020]. Disponible à l'adresse : <https://www.enssib.fr/bibliotheque-numerique/documents/66044-le-design-thinking-en-bibliotheque.pdf>

JACQUESSON, Alain et VON ROTEN, Gabrielle, 2019. *Histoire d'une (r)évolution : l'informatisation des bibliothèques genevoises, 1963-2018*. Genève : L'Esprit de la Lettre éditions. Bibliothéchos. ISBN 9782940587117.

JORF, 2014. Légifrance - Publications officielles - Journal officiel - JORF n° 0179 du 05/08/2014. *Légifrance.gouv.fr* [en ligne]. 5 août 2014. [Consulté le 13 juillet 2020]. Disponible à l'adresse : <https://beta.legifrance.gouv.fr/download/pdf?id=chzHL-IOL0uKBUDrMBLQw4tu-vb1CNBf2Gc37R1UyLU=>

LASCARIDES, Michael et VERSHBOW, Ben, 2014. What's on the Menu? : Crowdsourcing at the New York Public Library. In: RIDGE, Mia. *Crowdsourcing our Cultural Heritage*. Farnham: Ashgate, pp.113-137. ISBN 978-1-4724-1022-1.

LEON, Sharon M., 2014. Build, Analyse and Generalise: Community Transcription of the Papers of the War Department and the Development of *Scripto*. In: RIDGE, Mia. *Crowdsourcing our Cultural Heritage*. Farnham: Ashgate, pp.89-111. ISBN 9781472410221

MCKINLEY, Donelle, 2015. *Design principles for crowdsourcing cultural heritage*. [en ligne]. Wellington : Victoria University of Wellington. Rapport de thèse. Dernière révision en juillet 2015. [Consulté le 17 juin 2020]. Disponible à l'adresse : <http://nonprofitcrowd.org/wp-content/uploads/2015/03/McKinley-crowdsourcing-principles-July2015.pdf>

MOIREZ, Pauline, 2013. Bibliothèques, crowdsourcing, métadonnées sociales. *Bulletin des Bibliothèques de France (BBF)* [en ligne]. 2013. N° 5, pp. 32-36. [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://bbf.enssib.fr/consulter/bbf-2013-05-0032-007>

MOIREZ, Pauline, MOREUX, Jean Philippe et JOSSE, Isabelle, 2013. État de l'art en matière de crowdsourcing dans les bibliothèques numériques. *BnF* [en ligne]. [Consulté le 13 juin 2020] Disponible à l'adresse : https://www.bnf.fr/sites/default/files/2018-11/crowd_sourcing_etat_art.pdf

MOIREZ, Pauline, 2017. Construire et enrichir ensemble les savoirs : crowdsourcing et patrimoine numérisé. In : DUJOL, Lionel (dir.). *Communs du savoir et bibliothèques*. Paris : Editions du Cercle de la Librairie. Bibliothèques, pp. 129-139. ISBN 978-2-765-41530-5

MOYLE, Martin, TONRA, Justin et WALLACE, Valerie, 2011. Manuscript Transcription by Crowdsourcing: Transcribe Bentham. *LIBER Quarterly* [en ligne]. 11 février 2011. Vol. 20, n° 3-4, pp. 347-356. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://www.liberquarterly.eu/articles/10.18352/lq.7999/>

NÉROULIDIS, Ariane, 2015. *Le crowdsourcing appliqué aux archives numériques : concepts, pratiques et enjeux* [en ligne]. Lyon : ENSSIB. Mémoire de recherche. [Consulté le 13 juillet 2020]. Disponible à l'adresse :

<https://www.enssib.fr/bibliotheque-numerique/notices/66008-le-crowdsourcing-applique-aux-archives-numeriques-concepts-pratiques-et-enjeux>

OFFICE FÉDÉRALE DE LA CULTURE (OFC), 2019. Statistiques de poche de la culture en Suisse. *Confédération Suisse* [en ligne]. 27 septembre 2019. [Consulté le 15 mai 2020]. Disponible à l'adresse : <https://www.bak.admin.ch/bak/fr/home/actualites/archives-des-actualites/aktuelles-2019/neue-ausgabe-taschenstatistik-kultur-schweiz.html>

OFFICE FÉDÉRALE DE LA STATISTIQUE (OFS), 2020. Utilisation d'internet, données médicales et e-ressources des bibliothèques. *Confédération Suisse* [en ligne]. 21 juillet 2020. [Consulté le 3 août 2020]. Disponible à l'adresse :

<https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees/tableaux.assetdetail.13507402.html>

OOMEN, Johan et AROYO, Lora, 2011. Crowdsourcing in the cultural heritage domain: opportunities and challenges. In: *Proceedings of the 5th International Conference on Communities and Technologies* [en ligne]. Brisbane, Australia : Association for Computing Machinery. 29 juin 2011. pp. 138–149. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://dl.acm.org/doi/10.1145/2103354.2103373>

O'REILLY, Tim, 2005. What is Web 2.0 ? *O'Reilly* [en ligne]. 30 septembre 2005. [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>

OWENS, Trevor, 2014. Making Crowdsourcing Compatible with the Missions and Values of Cultural Heritage Organisations. In : RIDGE, Mia. *Crowdsourcing our Cultural Heritage*. Farnham. : Ashgate, pp.269-280. ISBN 978-1-4724-1022-1.

PHILIPPEAUX, Franck, 2015. Le Musée dauphinois et ses publics, quelles participations ? In : BATS, Raphaëlle (éd.). *Construire des pratiques participatives dans les bibliothèques*. Villeurbanne : Presses de l'enssib. La Boîte à outils, pp. ISBN 978-2-37546-069-6.

RIDGE, Mia, 2013. From Tagging to Theorizing: Deepening Engagement with Cultural Heritage through Crowdsourcing. *Curator: The Museum Journal* [en ligne]. 1 octobre 2013. Vol. 56, n° 4, pp. 435-450. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://onlinelibrary.wiley.com/doi/full/10.1111/cura.12046>

RIDGE, Mia, 2014. *Crowdsourcing our Cultural Heritage*. Farnham: Ashgate. ISBN 978-1-4724-1022-1.

RIJKSMUSEUM, 2020. @rijksmuseum. We love this Stay At Home Challenge! *Post Facebook* [en ligne]. 19 mars 2020, 14:33. [Consulté le 20 juillet 2020]. Disponible à l'adresse : <https://www.facebook.com/rijksmuseum/posts/10163090212705177>

SABATINI CHOQUARD, Laetizia, 2018. *Co-design, design thinking... : méthodes innovantes pour redessiner les services de l'Infothèque dans le cadre de son projet de réaménagement* [en ligne]. Genève : Haute école de gestion de Genève. Travail de Master. [Consulté le 8 avril 2020]. Disponible à l'adresse : <https://doc.rero.ch/record/323504?ln=fr>

SCHMIDT, Aaron et ETCHES, Amanda, 2016. *Utile, utilisable, désirable : Redessiner les bibliothèques pour leurs utilisateurs*. Nouvelle édition. Villeurbanne : Presses de l'enssib. ISBN : 978-2-3754-6053-5.

SEVERSON, Sarah et SAUVÉ, Jean-Sébastien, 2019. Crowding the Library : How and why Libraries are using Crowdsourcing to engage the Public. *Partnership : The Canadian Journal of Library and Information Practice and Research* [en ligne]. 15 juillet 2019. Vol. 14, n° 1. [Consulté le 24 juin 2020]. Disponible à l'adresse : <https://journal.lib.uouelph.ca/index.php/perj/article/view/4632>

TERRAS, Melissa, 2016. Crowdsourcing in the Digital Humanities. In : SCHREIBMAN, Susan, SIEMENS, Ray et UNSWORTH, John (éd.), *A New Companion to Digital Humanities!* [en ligne]. Oxford : Wiley-Blackwell, pp. 420-439. [Consulté le 17 juin 2020]. Disponible à l'adresse : <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118680596.html>

VIKHROVA, Anne, 2017. *L'évaluation de la méthode du crowdsourcing pour la transcription de manuscrits* [en ligne]. Grenoble : Université Grenoble Alpes. Thèse. [Consulté le 22 juin 2020]. Disponible à l'adresse : <https://tel.archives-ouvertes.fr/tel-01834993>

VILLE DE GENÈVE, 2020a. Chasse aux œufs au cœur des Collections de la Ville de Genève. *Ville de Genève* [en ligne]. 9 avril 2020, 23 :10. [Consulté le 15 mai 2020]. Disponible à l'adresse : <https://www.geneve.ch/fr/chasse-oeufs-virtuelle>

VILLE DE GENÈVE, 2020b. Positionner Genève comme ville numérique. *Ville de Genève* [en ligne]. 30 janvier 2020, 13 :10 [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://www.geneve.ch/fr/actualites/dossiers-information/geneve-ville-numerique/positionner>

VILLE DE GENÈVE 2020c. Département de la culture et de la transition numérique. *Ville de Genève* [en ligne]. 13 juillet 2020, 16:26. [Consulté le 14 juillet 2020]. Disponible à l'adresse : <https://www.geneve.ch/fr/autorites-administration/administration-municipale/departement-culture-transition-numerique>

VILLE DE GENÈVE, 2020d. A propos. *Mirabilia* [en ligne]. 16 juillet 2020. [Consulté le 16 juillet 2020]. Disponible à l'adresse : <https://mirabilia.ch/fr/a-propos>

WIEDERKEHR, Stefan, 2019. Open data for the crowd: an account of citizen science at ETH Library. *LIBER Quarterly* [en ligne]. 23 décembre 2019. Vol. 29, n° 1, pp. 1-10. [Consulté le 24 juin 2020]. Disponible à l'adresse : <http://www.liberquarterly.eu/articles/10.18352/lq.10294/>

Annexe 1 : Grille d'expérience

Nom (facultatif) :

Date :

Plateforme de transcription collaborative choisie :

	Première connexion à la plateforme	Création d'un compte <i>(facultatif)</i>	Choix d'un document à transcrire	Transcription du document	Autres étapes, préciser <i>(facultatif)</i>
Impression générale					
Actions, démarches à effectuer					
Questions que vous vous posez					
Bons moments, bonnes surprises, bonnes idées					
Irritants, motifs d'insatisfaction, de découragement					
Idées d'amélioration					

Commentaires :

Annexe 2 : Guides d'entretien

I. David Glaser – 16.04.2020

Autour des objectifs de NH.ch et de leur évolutions / fonctions / différence entre les souhaits et la réalité

1. **Objectifs** de départ de la plateforme ? Que sont-ils devenus au fil des années ?

Quels objectifs au départ ? A quoi sont-ils arrivés ? Ont-ils dévié des objectifs de base ?

(Quel était le public-cible visé au début par NH.ch, quel est-il aujourd'hui ? Selon leur récent sondage (<https://notrehistoire.ch/entries/04YzADkzWqK>), il s'agit à 73 % d'hommes et à 80 % par des 55 ans et plus, surtout dans le tertiaire (23 %) et l'enseignement (16 %))

Pourquoi les **refontes**, la deuxième arrivant assez vite (1ère : 2016 - 2e : 2019) ? Sur quels points ces refontes ont-elles porté (objectifs), et avec quel succès ? Sur quoi ont porté les critiques ?

2. Sur votre plateforme, **qu'est-ce qui marche bien** - par exemple, qu'en est-il de #ruesvideschallenge, la galerie que vous animez ? Et en quoi consiste l'animation d'une galerie ? Comment les mettre en valeur ? Y a-t-il un "nettoyage" ? = stratégies de valorisation ?

Pourquoi, selon lui ? **facteurs de succès** ?

Qu'en est-il du petit jeu (sympa d'ailleurs) Date my Romandie (<https://lab.notrehistoire.ch/exp1#/>) ? Comment y sont-ils arrivés ? Combien de temps le projet a-t-il pris ? Comment s'est fait le choix des images : des images souvent regardées, ou au contraire assez peu prisées, un mélange des deux ? Et comment l'outil a-t-il été testé ?

Cette idée d'un Lab est-elle poursuivie ? Et la collaboration avec l'Institut d'ingénierie des médias de la HEIG-VD (et qui au sein du MEI - Media Engineering Institute, <https://heig-vd.ch/rad/instituts/mei/> - voir Laurent Bolli, spécialiste "Utilisateurs et interfaces") ?

Et qu'en est-il de **L'Inédit** - un quotidien ? quel est son objectif / sont ses objectifs ?

3. Ce qui **n'a pas bien marché** ? Et pourquoi selon lui ?

Questions techniques (outils, maintenance, équipe, présence en temps ? EPT ? Budget ?

1. Technique : Comment se sont passés vos **choix techniques** : choix de l'outil (procédure de sélection) - quels ont été vos critères des choix ? Principales évolutions sur les deux refontes ?
2. Qu'en est-il de la **modération** ? comment les membres de la communauté sont **accompagnés** et comment sont-ils **récompensés** ?

Données chiffrées disponibles ? Par ex., état (combien de documents ?) et **mouvements** : **quelle progression d'année en année** ? types de données récoltées ?

(1127 membres actifs récemment - 5000 en tout environ, et bientôt 100'000 doc. !)

Comment gère-t-il la communauté (les diverses catégories) ? les nouveaux arrivants ?

Que sont les “membres d'excellence” ?

Que se passe-t-il avec les **membres dormants** ? Sont-ils effacés ? Sont-ils relancés ? **Mesures d'aide** pour les nouveaux arrivants et d'encouragement à la participation ?

3. A-t-on pensé à des **comités d'usagers** ?

Transcription collaborative de manuscrits / imprimés difficiles à océriser.

1. Qu'en est-il de la transcription collaborative de manuscrits ? Y a-t-il eu un projet ? Est-ce prévu ? S'il y en a eu, comment a-t-il été monté, comment cela a-t-il été communiqué ? Et avec qui ont-ils collaboré pour réfléchir à l'outil ?

Conclusion : ouvrir sur les partenariats

1. Quels sont vos partenariats, en dehors de la RTS ? Quels partenariats seraient possibles / envisageables avec la VdG ?

II. Barbara Roth – 04.05.2020

Présentation du projet, du contexte, de la méthodologie et des résultats attendus

Connaissance de projet de ce types et opinions

1. Projets participatifs en général : quelle expérience ? quelle opinion ?
2. Projets de transcription collaborative : quelle expérience ? quelle opinion ?

Expérience en tant qu'experte / conservatrice responsable Collections spéciales

1. Expérience et idées de valorisation des manuscrits en général
2. Est-ce qu'un tel projet a déjà été envisagé ? Si oui, quels fonds, quels publics, quel outil, sur quel modèle ?
3. Si elle devait ou avait dû mener un tel projet, à qui penser ? Quels risques ? Quels impératifs ? Quelles parties prenantes ? Quels partenariats ?

Ce futur projet

1. Son opinion en général : intérêt, craintes, curiosité, idées, suggestions
2. Les utilisateurs finaux : qui ? quel intérêt ? quelles réticences ?
3. Son avis sur : la modération, les interactions, les différents statuts des usagers
4. Intérêt pour ce projet et le suivre

III. Liliane Regamey – 25.05.2020

Présentation du projet, du contexte, de la méthodologie et des résultats attendus

Le projet e-newspaperarchives.ch

1. Origine du projet et stratégie
2. Publics cible, communication, notion de participation
3. Outils
4. Tutoriel
5. Modération et validation
6. Gestion des tags (vocabulaire libre ou contrôlé)
7. Intégration aux catalogues

Les résultats

1. Attentes, indicateurs, chiffres
2. Facteurs de succès
3. Facteurs d'échecs, de remise en question

IV. Alexa Renggli – 04.06.2020

Project, context, methodology, expected results

E-manuscripta

1. Origin and strategy
2. Public, target and communication
3. Tools
4. Tutorial and transcription rules
5. Moderation and validation
6. Integration in the catalogue

Results

1. Expectations, indicators, numbers
2. Successes
3. Failures, changes, and prospects

Annexe 3 : Questionnaire adressé aux responsables de la plateforme Mirabilia

À la demande du Département de la Culture et du Sport de la Ville de Genève (DCS), la filière Information documentaire de la Haute école de gestion mène en ce moment un mandat de recherche sur la *Transcription collaborative de manuscrits et imprimés : analyse de l'existant et recommandations pour les Musées de Genève*. En complément, Florence Burgy, étudiante du Master en Sciences de l'Information et assistante à la HEG, développe dans son Travail de Master un projet de mise en place d'un espace de transcription collaborative adapté à la plateforme Mirabilia.

Dans le but de correspondre au mieux aux objectifs du DCS et des institutions participantes, nous souhaitons vous soumettre le questionnaire suivant. Il comprend 6 questions et un tableau à remplir, et vise en premier lieu à mieux comprendre la vision de chacune des personnes impliquées dans ce projet. Il n'y a donc que des bonnes réponses !

Informations clés :

- Durée : 30-40 minutes
- Délai de renvoi : au plus tard le 15 mai 2020
- Pour questions et envoi : florence.burgy@hesge.ch

1. Avez- vous déjà fait l'expérience, dans un cadre professionnel ou privé, de **projets participatifs**, par exemple Muséomix, Biblioremix, *Crowdsourcing* ? Si oui, décrivez ce que vous avez aimé ou moins aimé, les succès et échecs, et ce que vous en reprenez.

2. Quel(s) **public(s)** pourraient, selon vous, participer à un projet de transcription collaborative de documents manuscrits issus des collections des Institutions de la Ville de Genève ?
 - a. Public(s) cible(s), i.e. ceux sur qui sera ciblée la médiation en priorité :

 - b. Public(s) potentiel(s), i.e. ceux que l'on pourrait également envisager :

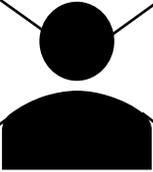
 - c. Usagers extrêmes, i.e. non-usagers, réfractaires, que l'on pourrait tenter de convaincre :

3. Choisissez l'un des **publics** que vous avez mentionnés, et essayez de vous mettre dans le rôle d'un usager de ce type. Décrivez, à votre avis :
 - Comment cet usager envisagerait un tel projet ?
 - Quelles pourraient être ses attentes, ses craintes ?
 - Qu'est-ce qui le motiverait ou le démotiverait ?
 - À quel(s) besoin(s) pourrait répondre ce projet ?

Vous pouvez répondre sous la forme que vous souhaitez (liste à puces, récit...), et/ou remplir la carte d'empathie à la page suivante (le personnage au centre étant l'utilisateur choisi).

Carte d'empathie (remplir les cases) :

1. L'utilisateur : qui est-il?	2. Ses objectifs, comment il saura qu'il les a atteints
6. Ce qu'il pense, ressent, ce qui le préoccupe	
5. Ce qui l'influence, ce que dit son entourage, à qui il fait confiance	3. Son environnement, les outils qu'il utilise
4. Ce qu'il fait, ce qu'il dit, son attitude en public	
7. Ce qu'il craint, ce qui le frustre, les obstacles et les risques sur sa route	8. Ses besoins, ses attentes, ce qui y répond



4. Quel **niveau de participation** pourriez-vous envisager pour un tel projet ? Merci de donner votre opinion sur les quelques exemples ci-dessous.



a) **Crowdsourcing simple** : les usagers transcrivent, des experts/professionnels vérifient et répondent aux questions, le comité éditorial prend les décisions relatives au choix de documents, à la gestion et maintenance de l'outil etc.

Oui Plutôt oui Plutôt non Non

Commentaires :



b) **Crowdsourcing plus libre** : les usagers transcrivent et peuvent aussi relire et valider les transcriptions d'autrui, répondre aux questions etc., le comité éditorial prend les décisions relatives au choix de documents, à la gestion et maintenance de l'outil etc.

Oui Plutôt oui Plutôt non Non

Commentaires :



c) **Consultation** : les usagers (ou un comité d'usagers) sont en outre consultés sur le choix de documents à transcrire (sur la base d'une liste transmise par les institutions participantes), sur l'évolution de l'outil de transcription collaborative etc, mais les décisions finales reviennent au comité éditorial.

Oui Plutôt oui Plutôt non Non

Commentaires :



d) **Conciliation** : les usagers (ou un comité d'usagers) choisissent en outre eux-mêmes les documents à transcrire (sur la base d'une liste transmise par les institutions participantes), proposent des améliorations de l'outil, des utilisations possibles des transcriptions une fois terminées (édition etc.), et les décisions sont prises en commun avec le comité éditorial.

Oui Plutôt oui Plutôt non Non

Commentaires :



e) **Pouvoir citoyen** : une fois le projet lancé, les décisions (documents à transcrire, outil, suites à donner au projet) reposent autant que possible sur les usagers (ou un comité d'usagers).

Oui Plutôt oui Plutôt non Non

Commentaires :

5. Avez-vous déjà des idées pour **exploiter et/ou valoriser les transcriptions produites** ?

6. Quelques **questions techniques** :

- a.** la plateforme et/ou le module de transcription collaborative devraient-ils être disponibles sur **mobile** ?

- b.** un « **réseau social** » dédié (*chat* et/ou forum pour les usagers) devrait-il être envisagé ?

- c.** la question de l'intégration des **publics non-voyants et mal-voyants** a-t-elle été considérée au niveau de la plateforme Mirabilia dans son ensemble, et si oui, de quelle manière ?

7. Enfin, afin de mieux comprendre **votre perception de ce projet**, nous souhaiterions que vous remplissiez en tout ou partie le tableau ci-dessous. Il s'agit de répondre selon votre avis, votre vision des choses – à nouveau il n'y a que des bonnes réponses !

Ce projet de transcription collaborative de documents manuscrits répondrait à...

	Pour ma hiérarchie	Pour mon département	Pour mes collaborateurs	Pour la plateforme Mirabilia	Pour le citoyen genevois lambda
Quels objectifs ... (stratégiques, socio-politiques, bibliothéconomiques, de recherche etc.)					
Quels besoins ...					
Quelles attentes ...					

Vous êtes arrivés au terme de ce questionnaire.

Je vous remercie pour votre précieuse collaboration.

Ce questionnaire est à renvoyer au plus tard le 15 mai 2020 à l'adresse suivante :

florence.burgy@hesge.ch

Genève, avril 2020

Annexe 4 : Questionnaire adressé au Comité éditorial de la plateforme Mirabilia

À la demande du Département de la Culture et du Sport de la Ville de Genève (DCS), la filière Information documentaire de la Haute école de gestion mène en ce moment un mandat sur la *Transcription collaborative de manuscrits et imprimés : analyse de l'existant et recommandations pour les Musées de Genève*. En complément, Florence Burgy, étudiante du Master en Sciences de l'Information et assistante à la HEG, développe dans son Travail de Master un projet de mise en place d'un espace de transcription collaborative adapté à la plateforme Mirabilia.

Dans le but de correspondre au mieux aux objectifs du DCS et des institutions participantes, nous souhaitons vous soumettre le questionnaire suivant. Il comprend 5 questions et un tableau à remplir, et vise en premier lieu à mieux comprendre la vision de chacune des personnes impliquées dans ce projet. Il n'y a donc que des bonnes réponses !

Informations clés :

- Durée : 30-40 minutes
- Délai de renvoi : au plus tard le 15 mai 2020
- Pour questions et envoi : florence.burgy@hesge.ch

1. Avez- vous déjà fait l'expérience, dans le cadre professionnel ou privé, de **projets participatifs**, par exemple Muséomix, Biblioremix, *Crowdsourcing* etc. ? Si oui, décrivez ce que vous avez aimé ou moins aimé, les succès et échecs, et ce que vous en reprenez.

2. Quel(s) **public(s)** pourraient, selon vous, participer à un projet de transcription collaborative de documents manuscrits issus des collections des Institutions de la Ville de Genève ?
 - a. Public(s) cible(s), i.e. ceux sur qui sera ciblée la médiation en priorité :

 - b. Public(s) potentiel(s), i.e. ceux que l'on pourrait également envisager :

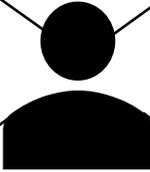
 - c. Usagers extrêmes, i.e. non-usagers, réfractaires, que l'on pourrait tenter de convaincre :

3. Choisissez l'un des **publics** que vous avez mentionnés, et essayez de vous mettre dans le rôle d'un usager de ce type. Décrivez, à votre avis :
 - Comment cet usager envisagerait un tel projet ?
 - Quelles pourraient être ses attentes, ses craintes ?
 - Qu'est-ce qui le motiverait ou le démotiverait ?
 - À quel(s) besoin(s) pourrait répondre ce projet ?

Vous pouvez répondre sous la forme que vous souhaitez (liste à puces, récit...), et/ou remplir la carte d'empathie à la page suivante (le personnage au centre étant l'usager choisi).

Carte d'empathie (remplir les cases) :

1. L'utilisateur : qui est-il?	2. Ses objectifs, comment il saura qu'il les a atteints
<p>6. Ce qu'il pense, ressent, ce qui le préoccupe</p>	
5. Ce qui l'influence, ce que dit son entourage, à qui il fait confiance	3. Son environnement, les outils qu'il utilise
<p>4. Ce qu'il fait, ce qu'il dit, son attitude en public</p>	
7. Ce qu'il craint, ce qui le frustre, les obstacles et les risques sur sa route	8. Ses besoins, ses attentes, ce qui y répond



4. Quel **niveau de participation** pourriez-vous envisager pour un tel projet ? Merci de donner votre opinion sur les quelques exemples ci-dessous.



a) **Crowdsourcing simple** : les usagers transcrivent, des experts/professionnels vérifient et répondent aux questions, le comité éditorial prend les décisions relatives au choix de documents, à la gestion et maintenance de l'outil etc.

Oui Plutôt oui Plutôt non Non

Commentaires :



b) **Crowdsourcing plus libre** : les usagers transcrivent et peuvent aussi relire et valider les transcriptions d'autrui, répondre aux questions etc., le comité éditorial prend les décisions relatives au choix de documents, à la gestion et maintenance de l'outil etc.

Oui Plutôt oui Plutôt non Non

Commentaires :



c) **Consultation** : les usagers (ou un comité d'usagers) sont en outre consultés sur le choix de documents à transcrire (sur la base d'une liste transmise par les institutions participantes), sur l'évolution de l'outil de transcription collaborative etc, mais les décisions finales reviennent au comité éditorial.

Oui Plutôt oui Plutôt non Non

Commentaires :



d) **Conciliation** : les usagers (ou un comité d'usagers) choisissent en outre eux-mêmes les documents à transcrire (sur la base d'une liste transmise par les institutions participantes), proposent des améliorations de l'outil, des utilisations possibles des transcriptions une fois terminées (édition etc.), et les décisions sont prises en commun avec le comité éditorial.

Oui Plutôt oui Plutôt non Non

Commentaires :



e) **Pouvoir citoyen** : une fois le projet lancé, les décisions (documents à transcrire, outil, suites à donner au projet) reposent autant que possible sur les usagers (ou un comité d'usagers).

Oui Plutôt oui Plutôt non Non

Commentaires :

- 5.** Quel(s) **fonds/document(s) manuscrit(s)** provenant de vos collections pourriez-vous proposer dès à présent pour un projet de transcription collaborative ?

Ce(s) document(s) doivent correspondre aux critères suivants :

- a.** Être en **français** (dans un premier temps)
- b.** Être doté(s) de **notices descriptives précises** (selon votre appréciation)
- c.** Être disponible(s) en une **version numérisée bien lisible**
- d.** Être **inédit(s)**
- e.** Ne pas être soumis à des **limitations de diffusion en ligne** (droit d'auteur etc.)
- f.** Être accessible(s) à un public profane quant au **type d'écriture manuscrite**
- g.** Capturer **l'attention du public** (que ce soit par le thème, l'auteur, le contexte, la typologie, etc.)

Fonds/documents suggérés :

6. Enfin, afin de mieux comprendre **votre perception de ce projet**, nous souhaiterions que vous remplissiez en tout ou partie le tableau ci-dessous. Il s'agit de répondre selon votre avis, votre vision des choses – à nouveau il n'y a que des bonnes réponses !

Ce projet de transcription collaborative de documents manuscrits répondrait à...

	Pour mon département/ ma hiérarchie	Pour mon institution	Pour les collections de mon institution	Pour les usagers de mon institution	Pour la plateforme Mirabilia	Pour le citoyen genevois lambda
Quels objectifs... (stratégiques, socio-politiques, bibliothéconomiques, de recherche etc.)						
Quels besoins...						
Quelles attentes...						

Vous êtes arrivés au terme de ce questionnaire.

Je vous remercie pour votre précieuse collaboration.

Ce questionnaire est à renvoyer au plus tard le 15 mai 2020 à l'adresse suivante :

florence.burgy@hesge.ch

Genève, avril 2020

Annexe 5 : Persona

Publics cibles, prioritaires

Profil A : Personne à la retraite



[Lien de l'image](#)

Nom : Isabelle A.

Âge : 68 ans

Activité : Retraitée

Centres d'intérêt : Art, littérature, culture, histoire

Appareils à disposition : ordinateur portable

User story : « Étant à la retraite, je fréquente régulièrement les institutions culturelles genevoises, et je souhaiterais consacrer une partie de mon temps à me rendre utile à ma communauté, tout en m'instruisant et me divertissant. »

Profil B : Amateur d'histoire, culture, littérature



[Lien de l'image](#)

Nom : Sylvain B.

Âge : 38 ans

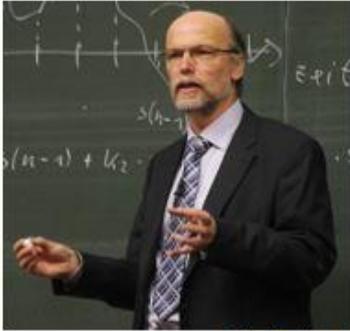
Activité : Community Manager

Centres d'intérêt : Photographie, littérature, histoire, paddle

Appareils à disposition : ordinateur portable, smartphone

User story : « Je n'ai pas beaucoup de temps à cause de mon travail et de ma vie de famille, mais j'aime bien l'Histoire et la littérature, et si je peux me rendre utile en m'instruisant de temps en temps je suis ravi. »

Profil C : Enseignant universitaire



[Lien de l'image](#)

Nom : Stéphane C.

Âge : 55 ans

Activité : Chargé d'enseignement en Histoire moderne

Centres d'intérêt : Histoire de Genève, art moderne, tennis

Appareils à disposition : ordinateur portable et de bureau, smartphone

User story : « J'aime pouvoir aider des amateurs d'Histoire à aborder des sources et à en comprendre le contenu et l'importance. En outre, j'aimerais que mes étudiants puissent s'exercer à la paléographie, afin d'être autonome dans leurs recherches et éditions de textes. »

Profil D : Enseignant au secondaire I ou II



[Lien de l'image](#)

Nom : Camille D.

Âge : 32 ans

Activité : Enseignante d'Histoire au cycle et au collège

Centres d'intérêt : Histoire, sciences, danse moderne

Appareils à disposition : ordinateur portable, smartphone

User story : « Je cherche toujours de nouvelles idées pour créer des séquences pédagogiques innovantes et interactives. De plus, je trouve essentiel que les élèves fassent l'expérience d'être confrontés à des sources brutes, pour mieux comprendre le travail des historiens. »

Profil E : Généalogiste



[Lien de l'image](#)

Nom : Philippe E.

Âge : 42 ans

Activité : Comptable

Centres d'intérêt : Généalogie, histoire locale, littérature suisse

Appareils à disposition : ordinateur portable, smartphone

User story : « Je consacre presque tout mon temps libre à des recherches en généalogie, et un accès direct aux documents qui me sont utiles, en ligne depuis chez moi, m'aiderait beaucoup dans mon travail, surtout s'ils sont déjà transcrits. »

Publics à convaincre

Profil F : Contributeur Wikipédia



[Lien de l'image](#)

Nom : Marine F.

Âge : 38 ans

Activité : Gestionnaire RH

Centres d'intérêt : Histoire des femmes, tricot, kickboxing

Appareils à disposition : ordinateur portable, smartphone

User story : « Dans mon temps libre, j'aime créer et corriger des articles Wikipédia. C'est une source très largement utilisée, et il est essentielle qu'elle soit à jour. Pour ce faire, j'ai besoin d'un accès direct et rapide aux sources qui m'intéresse. Je veux bien donner un peu de mon temps pour faciliter cet accès. »

Profil G : Étudiant Universitaire



[Lien de l'image](#)

Nom : Rachel G.

Âge : 25 ans

Activité : Etudiante en Master de Français

Centres d'intérêt : Littérature, art contemporain, natation

Appareils à disposition : ordinateur portable, smartphone

User story : « Pour un travail de séminaire, je dois m'appuyer sur la correspondance de l'auteur que j'étudie, et si elle est déjà transcrite, je pourrais faire des recherches plus rapidement. J'aimerais aussi pouvoir discuter de son œuvre avec d'autres personnes qui s'y intéressent. »

Profil H : Elève du secondaire II



[Lien de l'image](#)

Nom : Séverine H.

Âge : 17 ans

Activité : Élève de troisième année du collège

Centres d'intérêt : Lecture, théâtre, rock

Appareils à disposition : smartphone

User story : « Mon enseignante d'Histoire m'a demandé de réaliser un exposé sur une source genevoise numérisée. Je dois la transcrire moi-même et la commenter en classe. »

Profil I : Elève du secondaire I



[Lien de l'image](#)

Nom : Mathis I.

Âge : 13 ans

Activité : Elève du cycle d'orientation

Centres d'intérêt : Football, Mangas, TikTok

Appareils à disposition : smartphone

User story : « Ma prof nous a lancé le défi de transcrire les étiquettes de spécimens d'animaux. On ira ensuite au Museum pour voir si on reconnaît ces animaux dans les vitrines. »

Profil J : Chercheur universitaire



[Lien de l'image](#)

Nom : Christiane J.

Âge : 42 ans

Activité : Chercheuse en biologie végétale

Centres d'intérêt : Biologie, photographie, randonnée

Appareils à disposition : ordinateur portable et de bureau, smartphone

User story : « J'ai besoin de pouvoir faire des recherches dans des herbiers parfois anciens. Pouvoir faire une recherche par nom de plante dans les numérisations m'aiderait beaucoup. »

Public potentiel non prioritaire, usagers extrêmes

Profil K : Usager d'autres plateformes du même type



[Lien de l'image](#)

Nom : André K.

Âge : 64 ans

Activité : Libraire

Centres d'intérêt : Histoire, ornithologie, astronomie

Appareils à disposition : ordinateur portable

User story : « Je suis un adepte des plateformes de *crowdsourcing*. J'adore Europeana et Zooniverse. Je suis toujours à la recherche de nouveaux projets stimulants. »

Profil L : Gamer



[Lien de l'image](#)

Nom : Denise L.

Âge : 27 ans

Activité : Vendeuse dans un magasin d'électronique

Centres d'intérêt : Volley, guitare, jeux vidéo

Appareils à disposition : smartphone

User story : « Je passe beaucoup de temps sur mon smartphone. Quitte à faire ça, autant que ça soit parfois utile, que j'apprenne quelque chose au passage. »

Profil M : Bibliothécaire de prison



[Lien de l'image](#)

Nom : Claude M.

Âge : 37 ans

Activité : Bibliothécaire à Champ-Dollon

Centres d'intérêt : Musique classique, art de rue, jardinage

Appareils à disposition : ordinateur de bureau

User story : « Je souhaite contribuer à la réinsertion des détenus, en leur proposant des activités intellectuellement stimulantes et utiles à la société. »

Annexe 6 : Questionnaire adressé aux usagers potentiels – Enseignement et recherche en Histoire

À la demande du Département de la Culture et du Sport de la Ville de Genève (DCS), la filière Information documentaire de la Haute école de gestion mène en ce moment un mandat sur la *Transcription collaborative de manuscrits et imprimés : analyse de l'existant et recommandations pour les Musées de Genève*. En complément, je développe, dans mon Travail de Master en Sciences de l'Information, un projet de mise en place d'un espace de transcription collaborative adapté à la plateforme Mirabilia.

Dans le but de comprendre les opinions que des personnes évoluant dans les domaines de l'enseignement et de la recherche en Histoire pourraient avoir d'un tel projet, je souhaite vous soumettre le questionnaire suivant. Il comprend 4 questions et vise en premier lieu à mieux comprendre la vision de chacun. Il n'y a donc que des bonnes réponses !

Informations clés :

- Durée : 10-15 minutes
- Délai de renvoi : au plus tard le 18 mai 2020
- Pour questions et envoi : florence.burgy@hesge.ch

- 1.** Quel intérêt pourrait avoir, selon vous, une plateforme de transcription collaborative de manuscrits à Genève, et quel(s) usage(s) pourrai(en)t en être fait :
 - a.** Pour l'enseignement (Université, Secondaire I & II...) :
 - b.** Pour la recherche :
 - c.** Pour les loisirs, le temps libre (le vôtre ou celui d'autrui) :
- 2.** Avez-vous connaissance de besoins existants dans les trois domaines susmentionnés auxquels un tel projet pourrait répondre ?
- 3.** Si un tel projet devait voir le jour à Genève, quelles seraient vos attentes le concernant ? Vos espoirs ?
- 4.** Et vos craintes, vos doutes ?

Vous êtes arrivés au terme de ce questionnaire.

Je vous remercie pour votre précieuse collaboration.

Ce questionnaire est à renvoyer au plus tard le 18 mai 2020 à l'adresse suivante :

florence.burgy@hesge.ch

Genève, mai 2020

Annexe 7 : Questionnaire adressé aux usagers potentiels – Amateurs

À la demande du Département de la Culture et du Sport de la Ville de Genève (DCS), la filière Information documentaire de la Haute école de gestion mène en ce moment un mandat sur la *Transcription collaborative de manuscrits et imprimés : analyse de l'existant et recommandations pour les Musées de Genève*. En complément, je développe, dans mon Travail de Master en Sciences de l'Information, un projet de mise en place d'un espace de transcription collaborative adapté à la plateforme Mirabilia.

Dans le but de comprendre les opinions que des personnes amatrices de culture et d'Histoire locale et/ou globale pourraient avoir d'un tel projet, je souhaite vous soumettre le questionnaire suivant. Il comprend 3 grandes questions avec sous-questions, et vise en premier lieu à mieux comprendre la vision de chacun. Il n'y a donc que des bonnes réponses !

Informations clés :

- Durée : 15-20 minutes
- Délai de renvoi : au plus tard le 25 mai 2020
- Pour questions et envoi : florence.burgy@hesge.ch

- 1.** Concernant les bibliothèques et musées de la Ville de Genève⁹...
 - a.** Connaissez-vous ces institutions ? Si oui, lesquelles ? Et lesquelles avez-vous déjà fréquentées ?
 - b.** Expliquez en quelques mots pourquoi vous fréquentez ces institutions, ce qu'elles vous apportent, si vous y êtes attaché et pourquoi.
- 2.** Si une plateforme de transcription collaborative de manuscrits issus de ces institutions devait voir le jour à Genève, selon vous...
 - a.** Quel intérêt pourrait présenter ce projet ? Quel(s) usage(s) pourrai(en)t en être fait (par vous et/ou par d'autres) ?
 - b.** À quel(s) besoin(s) préexistant(s) ce projet pourrait-il répondre ?
 - c.** Quelles seraient vos attentes le concernant ? Vos espoirs ?
 - d.** Et vos craintes, vos doutes ?
- 3.** Quel(s) type(s) de document pourrai(en)t a priori vous intéresser ? Quel(s) thème(s) ? Quel(s) auteur(s) ?

⁹ Ariana, Archives de la Ville de Genève, Bibliothèque de Genève, Bibliothèques Municipales de la Ville de Genève, Conservatoire et Jardin Botaniques, Fonds d'Art Contemporain (MAMCO), Musée d'Art et d'Histoire, Musée d'Ethnographie de Genève, Musée d'Histoire Naturelle

Vous êtes arrivés au terme de ce questionnaire.

Je vous remercie pour votre précieuse collaboration.

Ce questionnaire est à renvoyer au plus tard le 25 mai 2020 à l'adresse suivante :

florence.burgy@hesge.ch

Genève, mai 2020

Annexe 8 : Questionnaire adressé aux usagers potentiels – Généalogistes

À la demande du Département de la Culture et du Sport de la Ville de Genève (DCS), la filière Information documentaire de la Haute école de gestion mène en ce moment un mandat sur la *Transcription collaborative de manuscrits et imprimés : analyse de l'existant et recommandations pour les Musées de Genève*. En complément, je développe, dans mon Travail de Master en Sciences de l'Information, un projet de mise en place d'un espace de transcription collaborative adapté à la plateforme Mirabilia.

Dans le but de comprendre les opinions que des généalogistes amateurs pourraient avoir d'un tel projet, je souhaite vous soumettre le questionnaire suivant. Il comprend 4 questions et vise en premier lieu à mieux comprendre la vision de chacun. Il n'y a donc que de bonnes réponses.

Informations clés :

- Durée : 10-15 minutes
- Délai de renvoi : au plus tard le 15 juin 2020
- Pour questions et envoi : florence.burgy@hesge.ch

- 1.** Quel intérêt pourrait avoir, selon vous, une plateforme de transcription collaborative de manuscrits à Genève, et quel(s) usage(s) pourrai(en)t en être fait :
 - a.** Pour votre activité de généalogiste amateur :

 - b.** Pour d'autres domaines (enseignement, recherche, loisirs...) :

- 2.** Avez-vous connaissance de besoins existants dans les domaines susmentionnés auxquels un tel projet pourrait répondre ?

- 3.** Si un tel projet devait voir le jour à Genève, quelles seraient vos attentes le concernant ? Vos espoirs ?

- 4.** Et vos craintes, vos doutes ?

Vous êtes arrivés au terme de ce questionnaire.

Je vous remercie pour votre précieuse collaboration.

Ce questionnaire est à renvoyer au plus tard le 15 juin 2020 à l'adresse suivante :

florence.burgy@hesge.ch

Genève, mai 2020

Annexe 9 : Questionnaire adressé aux usagers potentiels – Sciences naturelles

À la demande du Département de la Culture et du Sport de la Ville de Genève (DCS), la filière Information documentaire de la Haute école de gestion mène en ce moment un mandat sur la *Transcription collaborative de manuscrits et imprimés : analyse de l'existant et recommandations pour les Musées de Genève*. En complément, je développe, dans mon Travail de Master en Sciences de l'Information, un projet de mise en place d'un espace de transcription collaborative adapté à la plateforme Mirabilia.

Dans le but de comprendre les opinions que des personnes évoluant dans le domaine des Sciences naturelles pourraient avoir d'un tel projet, je souhaite vous soumettre le questionnaire suivant. Il comprend 4 questions et vise en premier lieu à mieux comprendre la vision de chacun. Il n'y a donc que de bonnes réponses.

Informations clés :

- Durée : 10-15 minutes
- Délai de renvoi : au plus tard le 15 juin 2020
- Pour questions et envoi : florence.burgy@hesge.ch

- 1.** Quel intérêt pourrait avoir, selon vous, une plateforme de transcription collaborative de textes manuscrits (ou autres types de *crowdsourcing*), à Genève, et quel(s) usage(s) pourrai(en)t en être fait :
 - a.** Pour l'enseignement :

 - b.** Pour la recherche :

 - c.** Pour les loisirs (les vôtres ou ceux d'autrui) :

- 2.** Avez-vous connaissance de besoins existants dans les domaines susmentionnés auxquels un tel projet pourrait répondre ?

- 3.** Si un tel projet devait voir le jour à Genève, quelles seraient vos attentes le concernant ? Vos espoirs ?

- 4.** Et vos craintes, vos doutes ?

- 5.** Quel(s) type(s) de documents devrai(en)t, à votre sens, être mis à disposition en priorité pour des projets de *crowdsourcing* ?

Vous êtes arrivés au terme de ce questionnaire.

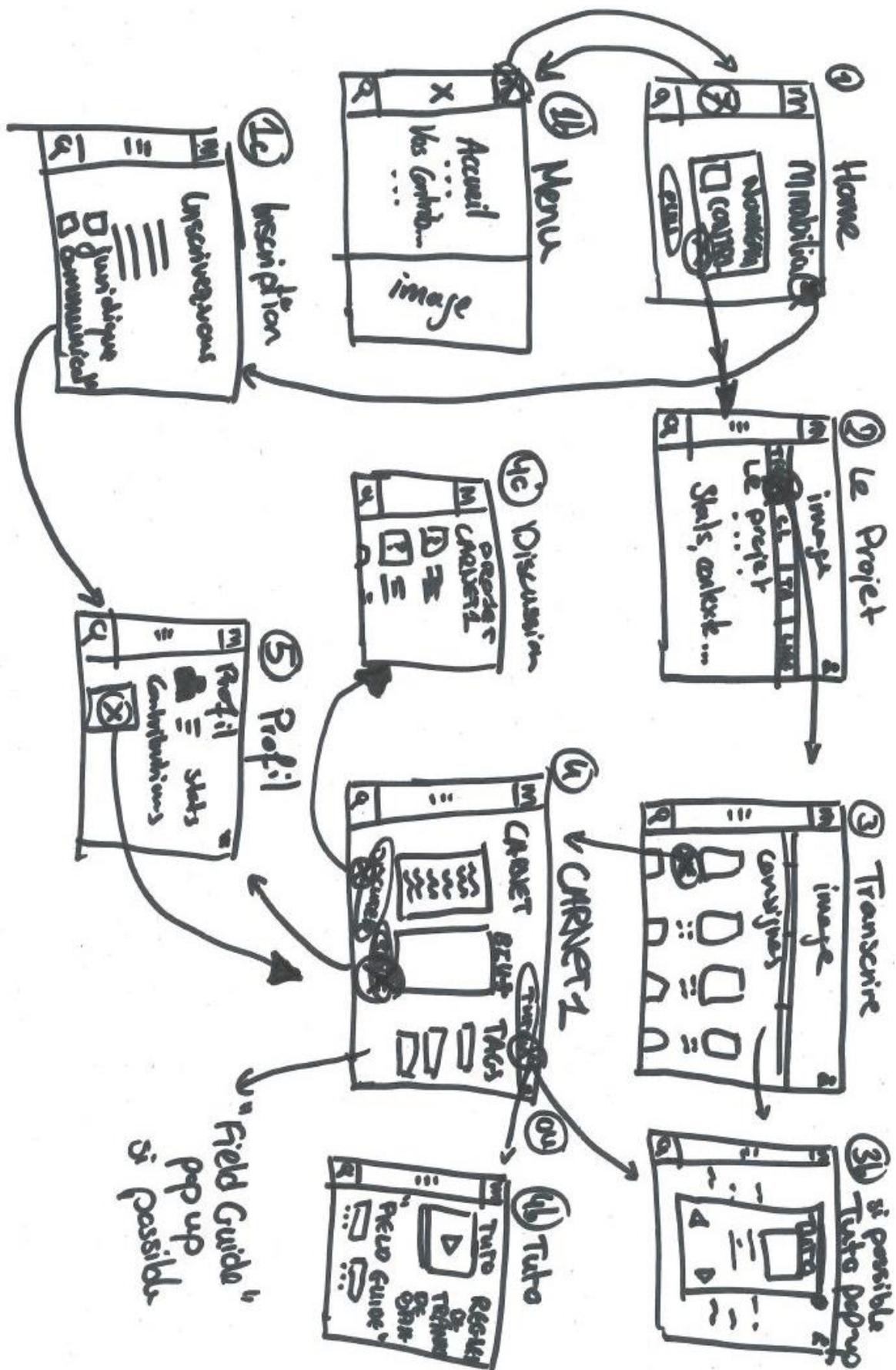
Je vous remercie pour votre précieuse collaboration.

Ce questionnaire est à renvoyer au plus tard le 15 juin 2020 à l'adresse suivante :

florence.burgy@hesge.ch

Genève, juin 2020

Annexe 10 : Prototype papier



Annexe 11 : Maquette

Page d'accueil

[Image](#)



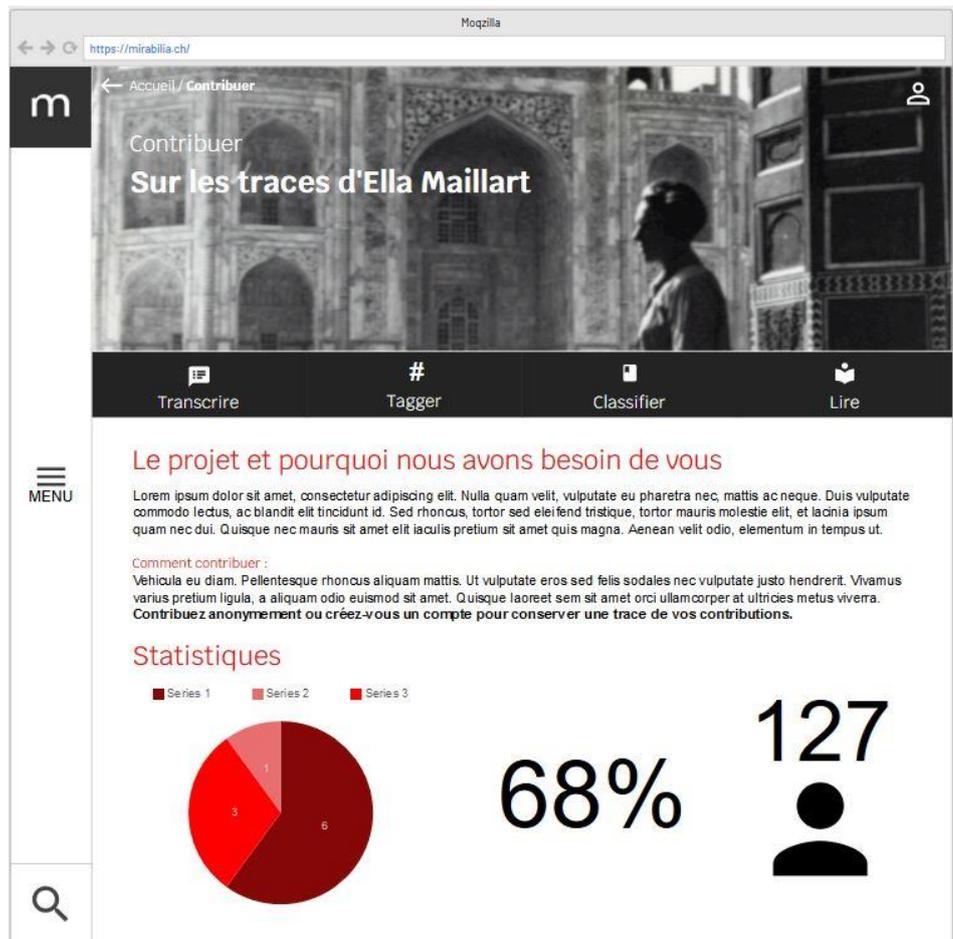
Menu

[Image](#)



Présentation du projet

[Image](#)



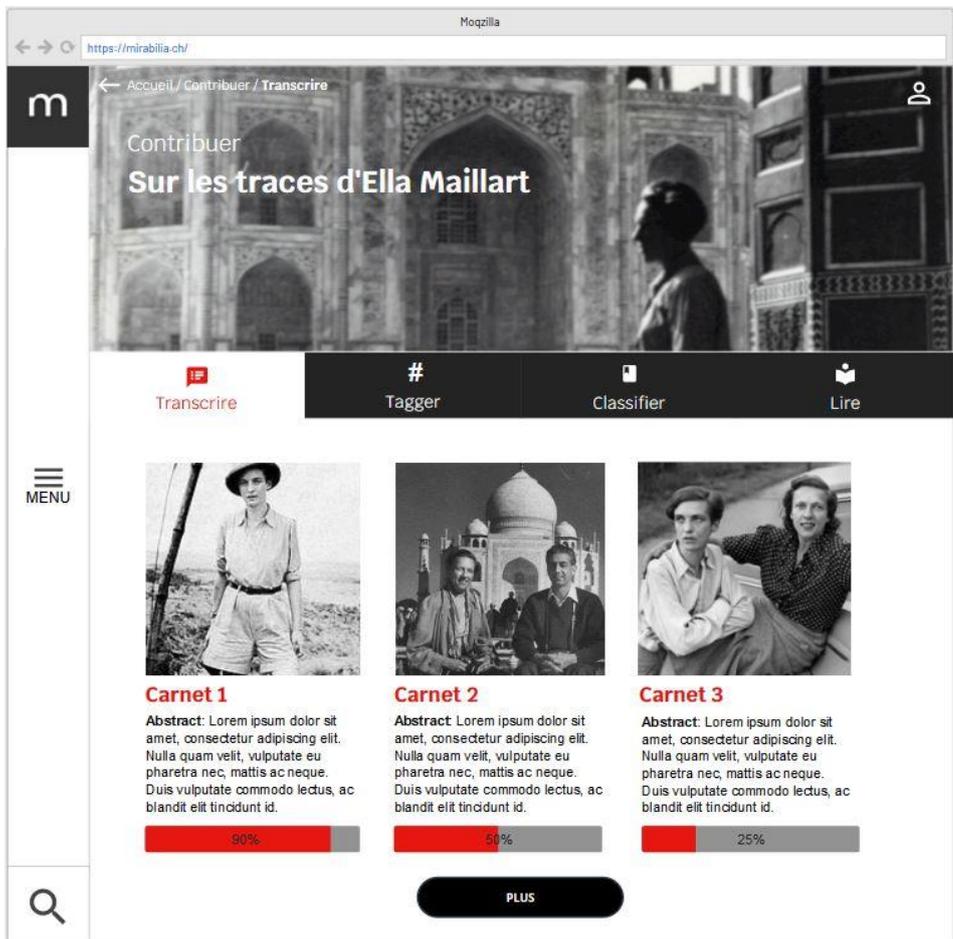
Transcrire

[Image bandeau](#)

[Image 1](#)

[Image 2](#)

[Image 3](#)



Profil

[Image](#)

mirabilia

Mon profil



Login : Lorem ipsum
E-mail : consectetur@adipisc.ch

Prénom (facultatif) :
Nom (facultatif) :
Date de naissance (facultatif) :
Profession (facultatif) :
Centres d'intérêt (facultatif) :

Mes statistiques



Mes contributions

Ella Maillart - Carnet 1



Annexe 12 : Maquette modifiée

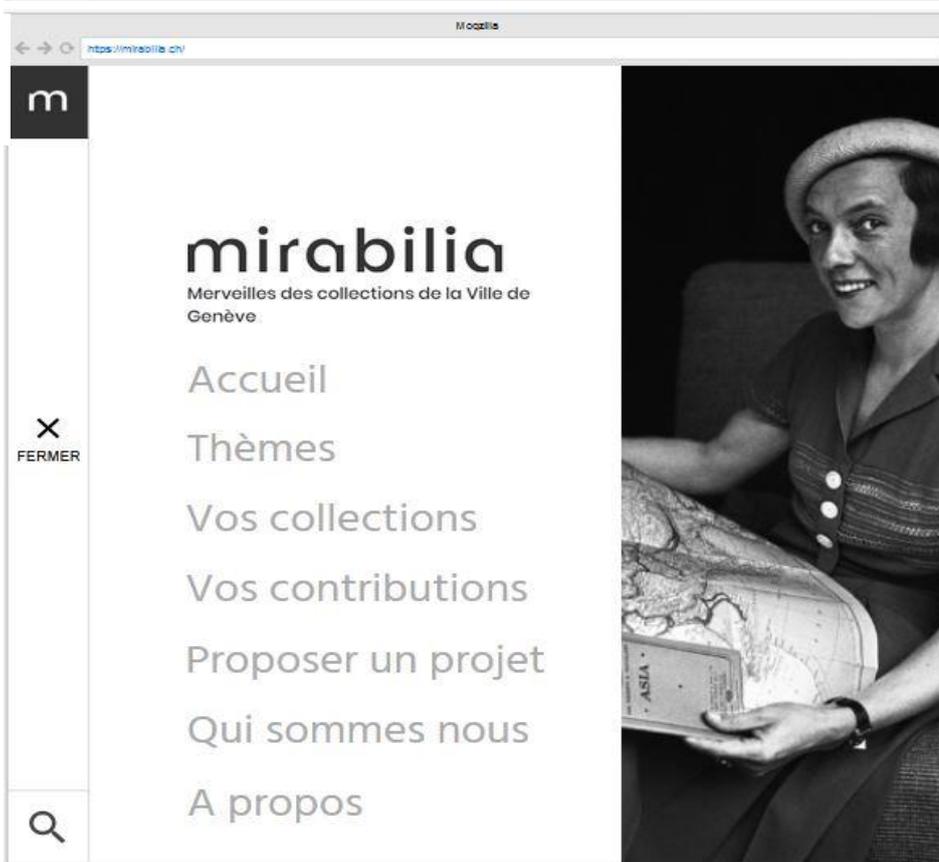
Page d'accueil

[Image](#)



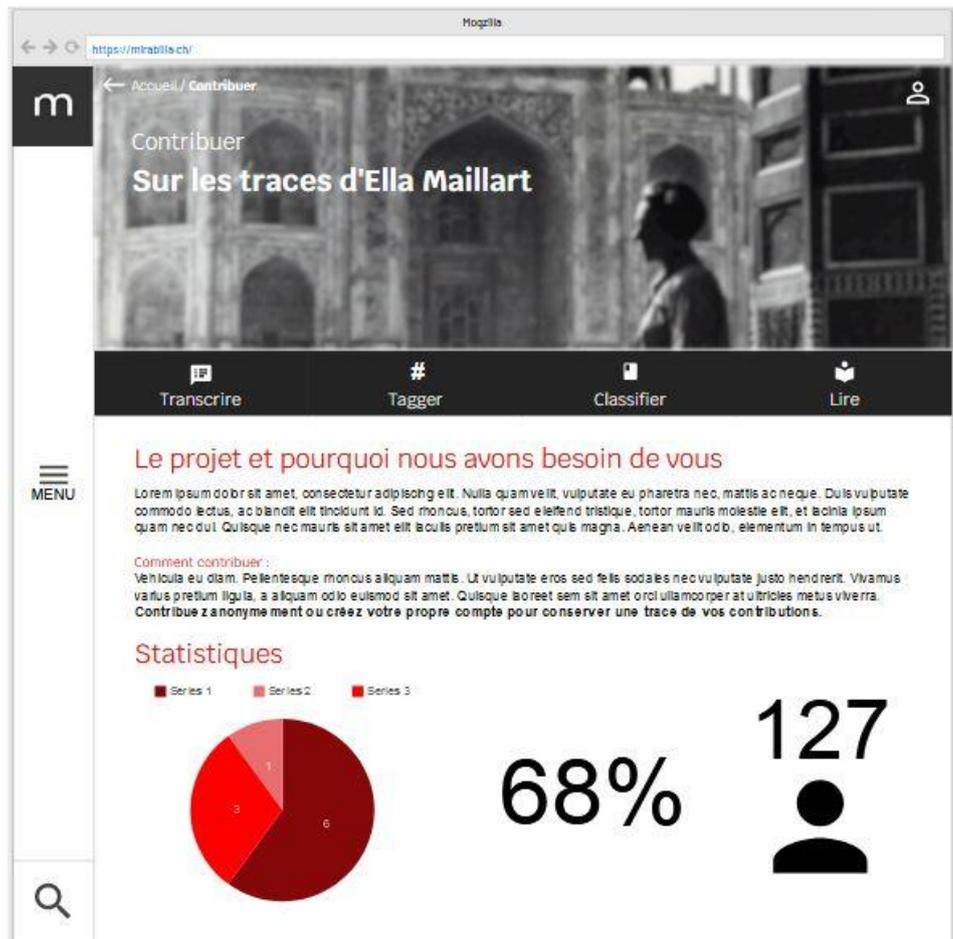
Menu

[Image](#)



Présentation du projet

[Image](#)



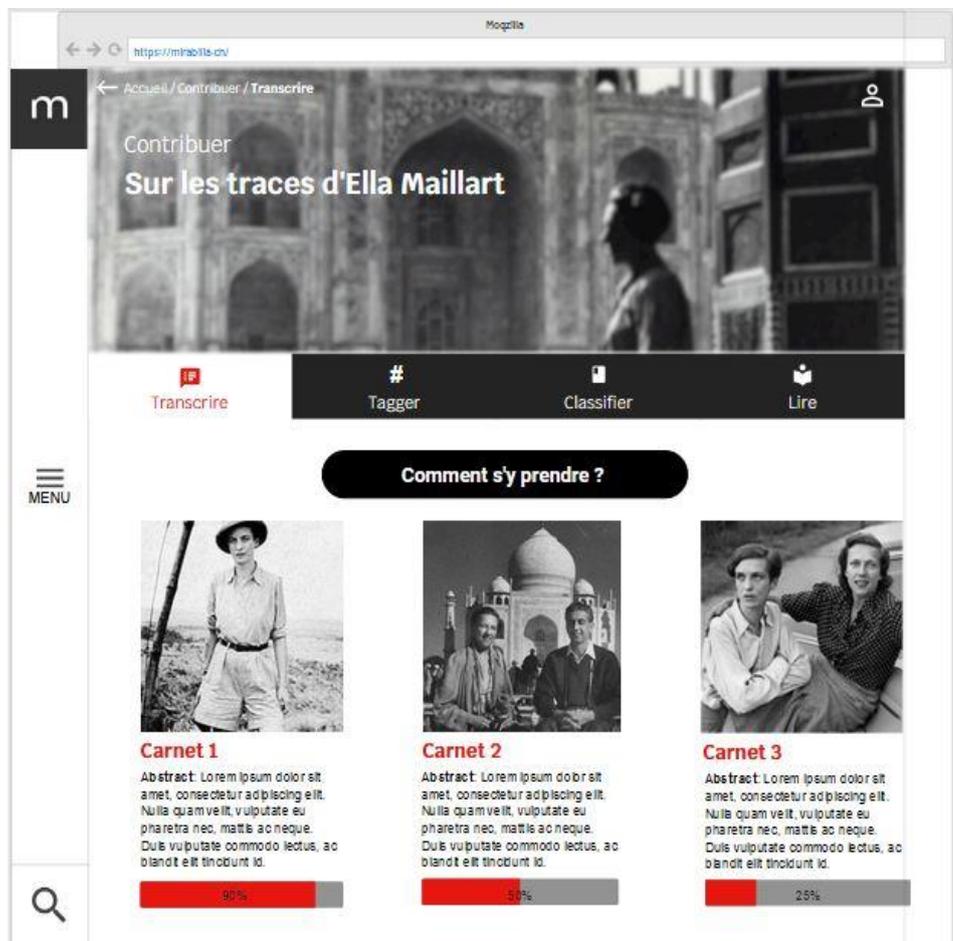
Transcrire

[Image bandeau](#)

[Image 1](#)

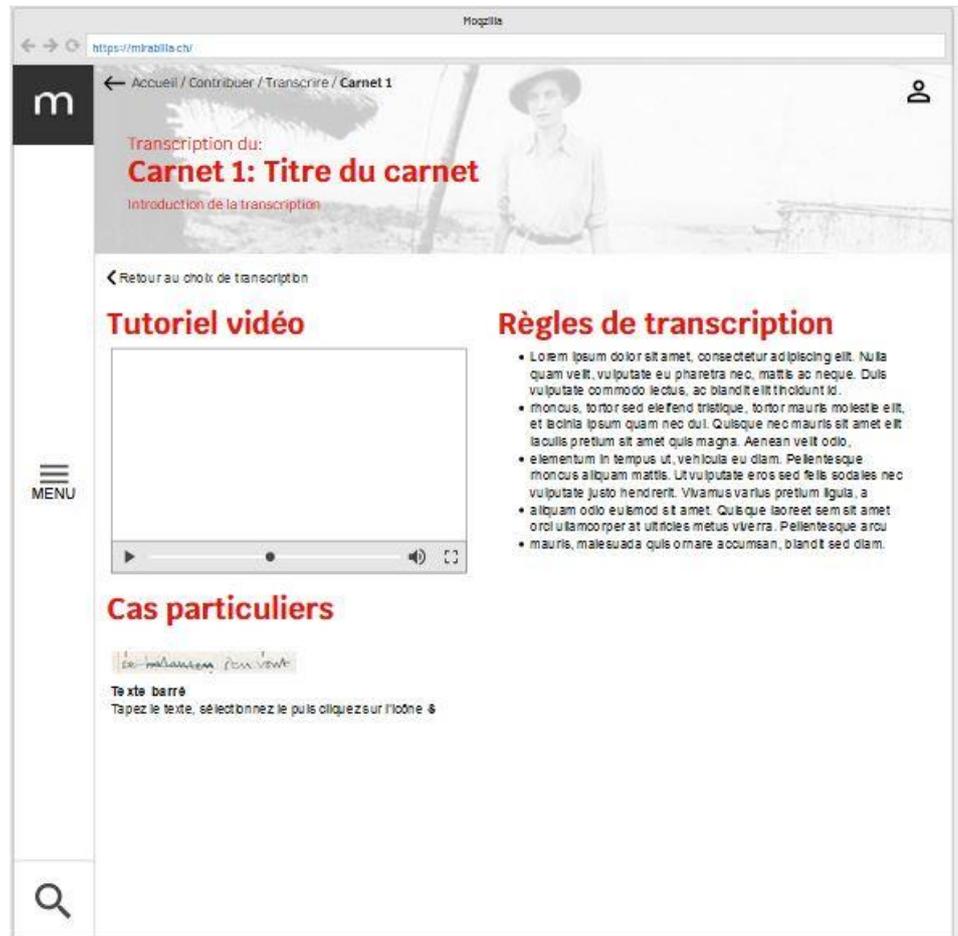
[Image 2](#)

[Image 3](#)



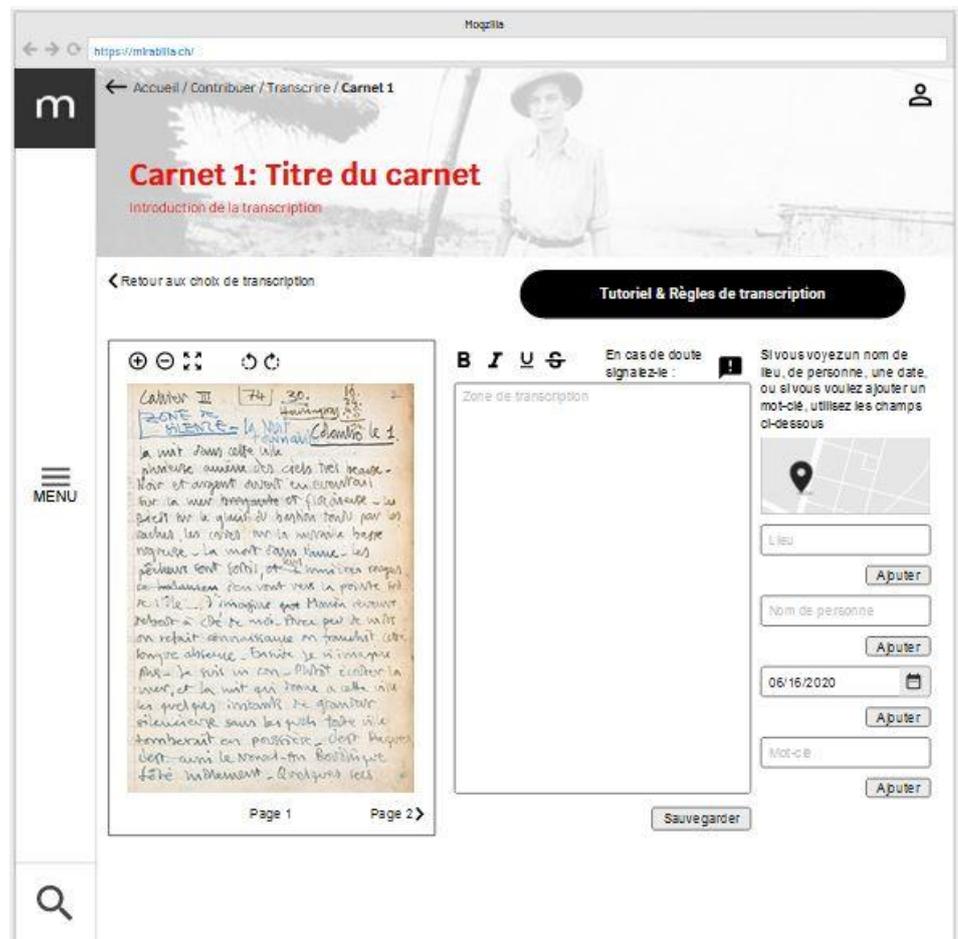
Carnet 1

[Image](#)



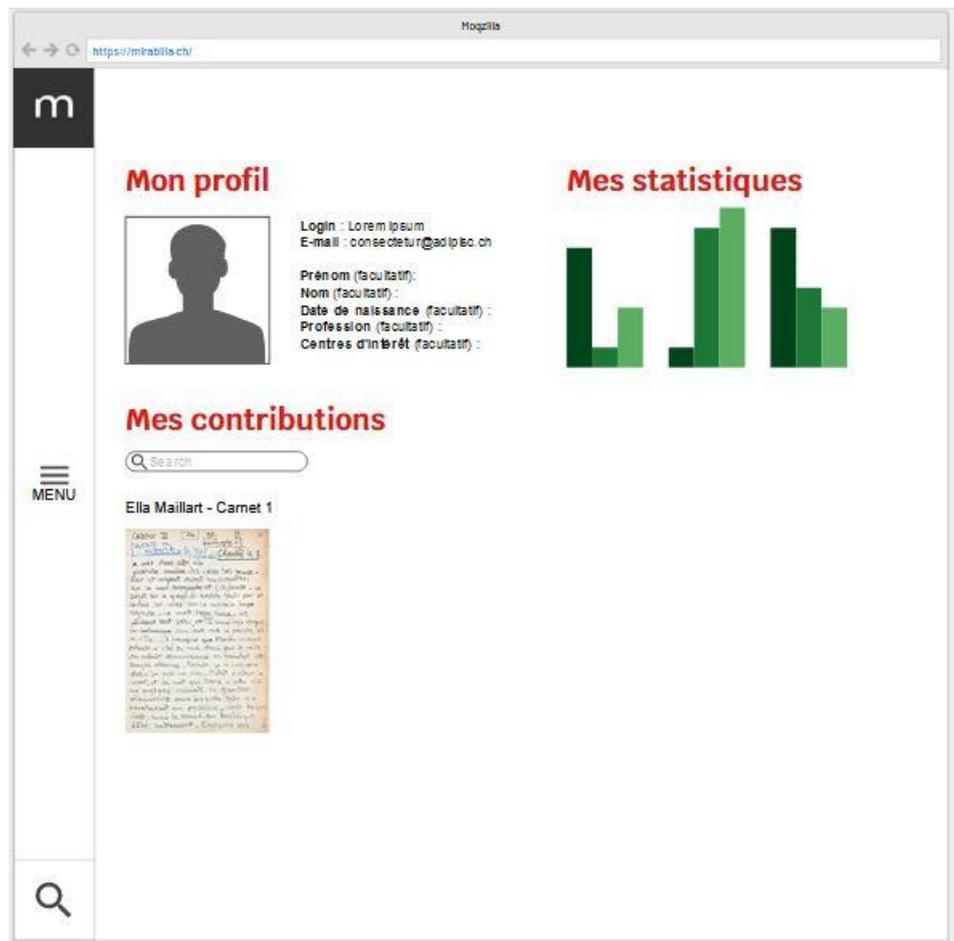
Tutoriel et règles de transcription

[Image bandeau](#) [Image cas](#)



Profil

[Image](#)



Annexe 13 : Outils de transcription collaborative – Google Sheet

	Institution		Hosted	Licence	Software	Text Type	TEI?	CMS Integration	Unique Features	URL	Code	
TextLab	John Bryant, et al, Hofstra University		No	?	?	Free-form	Yes	?	Direct annotation of TEI added/ del tags onto images.	http://mei.hofstra.edu/textlab.html		Melville Electronic Library
ScribeAPI	Zooniverse				Coffeescript / Ruby on Rails	Structured data + Free-form	No			https://github.com/zooniverse/scribeAPI	https://github.com/zooniverse/scribeAPI	
Wiki::Score (music scores in ABC)	?				DocuWiki				Collaborative music edition	http://www.wiki-score.org/soku.php		
National Archives Transcription Pilot Project	U.S. National Archives	1,000+ pages (300+ records)			Drupal	Free-Form		Drupal	Difficulty rating (Beginner, Intermediate, Advanced), lock out feature, commenting, links to online catalog	http://transcribe.archives.gov/		
Islandora TEI Editor (defunct)	UPEI, Robertson Library		No	GPL 3.0	Drupal/Fedora	Free-form	Yes	Fedora	TEI mark-up of documents hosted in Fedora	http://wiki.tei-c.org/index.php/IslandoraTEIEditor	https://github.com/Islandora/islandora_tei_editor	Public Records Office, Victoria http://prov.versl.edu.au/
World Archives Project	Anostry.com		Yes	Proprietary	Installed .exe client	Structured data (Genealogy)	?	?	difficulty rating, context based help, multiple archive sources	http://community.anostry.co.uk/awap		World Archive project and http://www.worldmemoryproject.org/ (essentially another "way in")
FieldData	Atlas of Living Australia/Gala Resources			Mozilla Public License 1.1	Java		No			http://www.ala.org.au/get-involved/citizen-science/fielddata-software/	http://code.google.com/p/ala-citizen-science/	http://volunteer.ala.org.au/project/index/42730
Intranda viewer	Intranda GmbH (http://www.digiverso.com)		optional	Proprietary	Java / Tomcat	Free-Form / Structured data	yes (as Export)	none		http://www.digiverso.com/digiproducts/viewer		http://is07.kobv.de/viewer/
Virtual Transcription Laboratory	Poznan Supercomputing and Networking Center		yes		Java Enterprise Edition + Tesseract	Structured data	No		Integrated OCR tool, trained for Polish historical documents also facility for OCR training is available. Nice and simple transcription editor which allows for text-image linking.	http://wit.synat.pozn.pl/		http://wit.synat.pozn.pl/wit-web/index.xhtml
T-PEN	St. Louis U Center for Digital Theology	As of 11/20/2012: 1,088 projects created by 538 users for a total of 61,162 lines transcribed	?	EPL 2.0	Java/Javascript	Line-based medieval	Yes	users can create export pipelines that can export transcriptions directly into a CMS database (such as Drupal)	Direct linking of transcription to lines of text in image	http://digital-editor.blogspot.com/	https://github.com/qinther/T-PEN	http://t-pen.org/TPEN/
Mirador (transcription branch)	IIF		No	Apache 2.0	JavaScript	Free-form	No		IIF	projectmirador.org	https://github.com/IIF/mirador/tree/transcription	
Alto Editor (kinda working, but obsolete)	IMPACT Centre of Competence		No	Apache 2.0	JavaScript/Ruby		No	none			https://github.com/ImpactCentre/alto-editor	
Scribe (Zooniverse)	Zooniverse		Upon application.	MIT	jQuery/Ruby on Rails	Structured data	No	none	Blind triple-keying, data linked to images		http://github.com/zooniverse/scribe	What's the Score at the Bodleian (earlier versions at OldWeather.org) FreeREG/FreeCEN rewrite for FreeUKGen will extend Scribe in 2012.
Bentham Transcription Desk	University of London Computer Centre; UCL Bentham Project	As of 31 August 2012: 4,168 manuscripts transcribed or partially-transcribed (c.2 million words, plus extensive TEI markup), of which 3,916 (93%) are complete	Yes	GPL2.0	MediaWiki		Yes		Full TEI mark-up support; customised toolbar to automatically apply TEI tags to transcriptions	http://www.ucl.ac.uk/transcribe-bentham	https://github.com/onothm/agen/sbp-transcription-desk	Transcribe Bentham; Public Record Office of Victoria's (Melbourne) Transcription Pilot
Wikisource	Wikimedia	http://toolsserver.org/~phe/statistics.php	Yes	GPL 2.0	MediaWiki	Free-form	No	Archive.org	Workflow management	http://en.wikisource.org/wiki/Main_Page	http://www.mediawiki.org/wiki/MediaWiki	NARA Citizen Archivist Dashboard
Dromio	Folger Library		Yes	?	PHP	Free-Form	Yes		Customizable set of elements, multiple transcriptions allowed, transcriptions can be collated	dromio.folger.edu http://collation.folger.edu/2014/12/a-transcribe-what/		

Scripto	Center for History and New Media at George Mason University		No	GPL 3.0	PHP library, MediaWiki	Free-form, wikitext	No	Omeka, WordPress, Drupal	Can be integrated into potentially any CMS or personal archive	http://scripto.org	https://github.com/chnm/scripto https://github.com/omeka/plugin-scripto https://github.com/chnm/scripto-wordpress-plugin https://github.com/chnm/scripto-drupal-module	Papers of the War Department, 1784-1800
Unbindery	Ben Crowder	10,000+ book pages transcribed (Project Gutenberg Thailand); 3000+ pages transcribed (Mormon Texts Project)	Yes	MIT	PHP/JavaScript	Free-form	No	is a CMS		http://bencrowder.net/coding/unbindery/	https://github.com/bencrowder/unbindery	Project Gutenberg Thailand http://gutenbergthai.org ; Mormon Texts Project http://mormontextsproject.org/
Biodiversity Volunteer Portal (rebranded as DigVol, below)	Atlas of Living Australia and Museum Australia	http://volunteer.ala.org.au/about/index	Yes	Mozilla Public License 1.1	Postgres/Java/Grails/Apache	?	?	?			https://code.google.com/p/ala-volunteer/source/checkout	Atlas of Living Australia
PyBOSSA	Citizen Cyberscience Centre/OKFN		?	AGPL 3.0	Python/GDocs	Tabular	No		Data entry via GDoc spreadsheet	http://pybossa.com/	https://github.com/PyBossa/pybossa	Transcribe Bleek & Lloyd
Scott Transcription	@anthonygoddard / zerosixzero.org		Yes	Apache 2	Ruby	Free-form			designed for Scott / Terra Nova log entries	http://scott-transcription.zerosixzero.org/	td	http://scott-transcription.zerosixzero.org/
FromThePage	Ben Brumfield / Brumfield Labs	7500+ pages transcribed, 2000+ indexed with 5000 subjects mentioned 25000 times as of 2017-07-01 (number from fromthepage.com, more pages on other sites.)	Yes	AGPL 3.0 (inquire for dual license)	Ruby on Rails	Free-form	Yes (as export)	Archive.org, Omeka, any IIIF-compliant digital library system	Semantic mark-up for indexing/annotation, OCR correction, collaborative translation	http://fromthepage.com/	http://github.com/benbrumfield/fromthepage/wiki	LA County Public Libraries, Fordham University, University of Texas, Indianapolis Public Library, Yaquina Head Lighthouses, New York Botanical Garden, University of Virginia Law Library, University College Dublin, San Diego Natural History Museum; Laurence M. Klauber Field Notes; Southwestern University; Zenas Matthews Diary; Rhodes College; Shelby Foote Diaries (private); Northwestern University Library; private project; Penn State U.; Phillip K. Dick folders; Mosman Council (NSW); WW1 letters, diaries, and inscriptions; U. Delaware: Civil War diary of Joseph Brown
Son of Suda On-Line	Integrating Digital Papyrology	As of 3/28/12: 3,694 submissions approved	No	GPL 3.0	Ruby on Rails	Free-Form	Yes	none	Git backend		https://github.com/sosol/sosol	papyfl.info
Persels	Tufts		Yes	GPL 3.0	Ruby on Rails	Free-Form	Yes			http://persels.org/	https://github.com/persels-project/persels_docs	
What's On the Menu? (custom Ruby on Rails app)	New York Public Library	796,136 dishes from 12,541 menus			Ruby on Rails	Structured data	No			http://menus.nypl.org/		
Transcribable	Propublica				Ruby on Rails (extension to ActiveRecord)	Structured	No	Docucloud		https://projects.propublica.org/tree-the-files/	https://github.com/propublica/transcribable	
Boolean Transcripitor	University College Cork, rep. of Ireland	377 items, 2495 scans	No	To Be Decided	Ruby on Rails+plugins, Bootstrap, jQuery+plugins	Plain text at the moment; in time transitioning to Markdown	Not Yet :)	No	Nice navigation, highlighting, YAML driven	http://boole-papers.electropolesis.org/	https://github.com/igtavlov/si-boolean-transcripitor	
DIYHistoryTranscribe	University of Iowa Libraries		Yes	?	Scripto, Omeka, Mediawiki	Free-form				http://diyhistory.lib.uiowa.edu/code.html	https://github.com/ui-libraries	http://diyhistory.lib.uiowa.edu/
Veridian Software	DL Consulting		Optional	Proprietary	Veridian	Allows free-form transcription as well as structured line-by-line transcription (if METS/ALTO is used for underlying data).	TEI is supported but METS/ALTO is recommended.	No	* Supports line-by-line correction of errors in OCR'd material (similar to the Trove and British Newspaper Archive sites). * Supports transcription of unstructured text (with side-by-side image/transcription display). * Supports structured metadata entry/transcription, in addition to text transcription. * Transcriptions/corrections instantly become "live" (i.e. are searchable).	http://veridiansoftware.com		http://cdnc.uor.edu/ http://cambridge.diconsulting.com/ http://virginiachronicle.com/

Letters of 1916	Trinity College Dublin, rep. of Ireland					Wordpress, Scripto, DIYHistory, Bentham TEI Toolbar						http://dn.tod.ie/letters1916/	
Itinerova-Editor	Stadsarchief Leuven / CCoH	4 000 records		GPL3	XRX/Javascript	Line based medieval	Yes (basic transcription features)				http://www.mom-wiki.uni-koeln.de/	https://subversion.rz.uni-koeln.de/trac/xist-A/browser/trunk/my/XRX/www/	http://www.itinerova.be/
VdU-Editor	Monasterium.net/HKI Cologne			GPL3	XRX/Javascript	Free-Form/structured data	possible (configurable for any xsd-schema)		full XML-editor with hidden XML-syntax		https://github.com/icauseu/mom-ca/wiki/How-to-Use-EditMOM3-Environment	https://github.com/icauseu/mom-ca	www.monasterium.net; Virtuelles deutsches Urkundennetzwerk: www.vdu.uni-koeln.de
Crowd-Ed (this tool is not precisely for transcription; it's designed for metadata editing which will, among other descriptions, indicate type vs. manuscript doc to help determine which documents require OCR and which require manual transcription)	Martha Berry Digital Archive Project	c. 1500 (edited, but not yet transcribed)	yes	free, open source	Zend / PHP	structured	no	Omeka	crowdsourced metadata editing aligned with Dublin Core		https://github.com/gsbodin/crowd-ed https://mbda.berry.edu		
Bentham TSX	UCL and transcriptorium		Yes				Yes				http://www.transcribe-bentham.da.ucl.ac.uk/TSX/ http://transcriptorium.eu/		
Canadian Census - 1901, 1911, etc	Automated Genealogy	>13 million census records transcribed, proofreading underway	Yes			Structured data (census records)	No				http://automatedgenealogy.com/census/		
http://pictureae.com/nl/	http://pictureae.com/nl/	as of 6/18/12: >240,000 records	yes	Closed		Structured data			Volunteers earn points that can be used to purchase scans		http://www.valehanden.nl		militeregisters.nl
Aplary (defunct)	Botanical Research Institute of Texas			?		Structured	No		OCR integration; Regions of Interest		http://www.aplaryproject.org/	https://github.com/jcast/Aplary	
Civil War Diaries & Letters Transcription Project (see DIYHistoryTranscribe)	The University of Iowa Libraries	As of 2/24/12: 9,043 pages				Free-Form					http://papyri.github.com/documentation/		
Digivol	Australian Museum												
Family Search Indexing	Family Search			Proprietary		Structured data (Genealogy)	No				https://indexing.familysearch.org/newuser/nuhome.jsf?_af=3.9.6		
Harold "Doc" Edgerton Project	MIT?					Free-Form					http://edgerton-digital-collections.org/notebooks		
North American Bird Phenology Program	USGS	560,271 cards transcribed; 1,104,494 cards scanned				Structured data	No				http://www.pwrc.usgs.gov/bp/		
Smithsonian Digital Volunteer	Smithsonian museum	Many projects completed					No				https://transcription.si.edu/		
typewriter	18c connect (Texas A&M)					Line-based OCR correction	No	ECCO/EEBO			http://www.18thconnect.org/typewriter/		
Hive	New York Times		No	Apache 2	Go, JSON, Elasticsearch	Free-form transcription and structured tags					http://nytlabs.com/blog/2014/12/09/hive-open-source-crowdsourcing-framework/	https://github.com/nytlabs/hive	Used for NYT's Madison project: http://madison.nytimes.com/

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Tool	Works (has working demo online or can be installed locally without needing to update code)	Open source, code available	Code last updated	Side-by-side image and text	Markup, tagging	XML	TEI authoring	TEI export	WYSIWYG transcription	Preview of marked up transcription	Adjustable based on transcriber's expertise	User can customize interface	Authentication	OAuth	Discussion attached to transcription	Multiple transcriptions allowed	Crowdsourced translation of transcribed content allowed
11	FromThePage	X	X	2017	X	X			X		X		X	X		X		X
12	Mirador (transcription)	X	X	2015	X	X											X	
13	Perseids (SOSOL)	X	X	2015	X	X	X	X	X		X			X	X	X		
14	PyBOSSA	X	X	2015	X									X	X			
15	ScribeAPI	X	X	2015	X	X						?		X	X		X	
16	Harold "Doc" Edgerton Project (excluded)																	
17	intrantra viewer (excluded)																	
18	Islandora TEI Editor (excluded)																	
19	Itineranovra-Editor (excluded)																	
20	Son of Suda On-Line (SOSOL)	X	X	2015		X	X	X	X		X			X	X	X		
21	Trascribable	X	X	2015	X	X								X	X			
22	Bentham TSX	X	?	?	X	X		X	?		X			X				
23	Biodiversity Volunteer Portal	X	?	?	X									X	X			
24	OpenScribe (excluded)																	
25	Letters of 1916 (Scripto)	X	?	?	X	X		X	?					X				
26	Canadian Census - 1901, 1911, etc	X			X									X				
27	Civil War Diaries & Letters Transcription Project (see DIYHistoryTranscribe)																	
28	Crowd-Ed (excluded)																	
29	Dromio (Folger library)	X			X	X	X	X	X	X	X		X	X			X	
30	Family Search Indexing (excluded)																	
31	FieldData (excluded)																	
32	National Archives Transcription Pilot Project		?															
34	North American Bird Phenology Program	X			X	X								X			?	
35	Scott Transcription	X			X	X											X	
36	Smithsonian Digital Volunteer typewriter	X			X	X								X		X		
37	VdU-Editor	X			X	X								X				
38	Velehanden.nl	X			X	X								X	X			
39	Veridian Software	X			X									X				
40	Virtual Transcription Laboratory	X			X									X				
41	What's on the menu (excluded)																	
42	Wiki:Score (music scores in ABC) (excluded)																	
43	Wikisource	X			X	X					X		X	X		X		
44	World Archives Project (excluded)																	

	A	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
1	Tool	Comparison of multiple transcriptions	IIIF integration	Image annotation	Text annotation	Tablet friendly	Smartphone friendly	OCR support	Manual image/text alignment	Semiautomati-c image/text alignment	Semi-automatic markup	CMS integration	Metadata editing	Versioning	Linked Open Data integration	Clear guidelines	Context-based help (e.g. not just separated guidelines)	Editorial workflow (some users can revise / approve transcription)	User groups	Extendable with plug-ins	Support available (e.g. email, chat, tickets)
11	FromThePage		X			X		X			X	X		X	?	X	?	X			X
12	Mirador (transcription)			X		tbd	tbd		X						X						
13	Perseids (SOSOL)			X		tbd	tbd		X		X		X			X		X			
14	PyBOSSA					tbd	tbd				X		X	X							
15	ScribeAPI			X					X								X			X	
16	Harold "Doc" Edgerton Project (excluded)																				
17	Intranda viewer (excluded)																				
18	Islandora TEI Editor (excluded)																				
19	Itineranovae-Editor (excluded)																				
20	Son of Suda On-Line (SOSOL)					tbd	tbd				X		X	X		X					X
21	Transcribable					tbd	tbd										X				
22	Bentham TSX					?	?			X						X		X			
23	Biodiversity Volunteer Portal												X			X	X	X			
24	OpenScribe (excluded)																				
25	Letters of 1916 (Scripto)					tbd	tbd							X		X					
26	Canadian Census - 1901, 1911, etc					tbd	tbd											X			
27	Civil War Diaries & Letters Transcription Project (see DIYHistoryTranscribe)																				
28	Crowd-Ed (excluded)																				
29	Dromio (Folger library)	X				tbd	tbd						X				X	X			
30	Family Search Indexing (excluded)																				
31	FieldData (excluded)																				
32	National Archives Transcription Pilot Project																				
34	North American Bird Phenology Program					tbd	tbd									X					
35	Scott Transcription					tbd	tbd														
36	Smithsonian Digital Volunteer					tbd	tbd									X		X			
37	typewright							X	X	X	X					X					
38	Vdu-Editor																				
39	Velehanden.nl					tbd	tbd			X											
40	Veridian Software					tbd	tbd	X		X											
41	Virtual Transcription Laboratory			X		tbd	tbd	X	X	X											
42	What's on the menu (excluded)																				
43	Wiki: Score (music scores in ABC) (excluded)																				
44	Wikisource					tbd	tbd						X	X		X	X				
45	World Archives Project (excluded)																				