Le *crowdsourcing* dans les bibliothèques

Mathieu Andro

Institut national de la recherche agronomique (INRA),
Délégation à l'Information scientifique et technique (DIST)

Résumé : L'externalisation auprès de foules d'internautes a des origines anciennes. Aujourd'hui le *crowdsourcing* peut être bénévole ou rémunéré, volontaire ou involontaire, prendre la forme de jeux. Il peut faire appel à des types de motivations très diverses. Dans les bibliothèques, le *crowdsourcing* n'est pas sans conséquences sur le mode de fonctionnement. Il génère de nombreuses critiques et est parfois qualifié d'ubérisation des bibliothèques.

Mots-clés : bibliothèques numériques, numérisation, *crowdsourcing*, crowdfunding, gamification

Abstract : Outsourcing to crowds has ancient origins. Today, crowdsourcing can be free or paid, voluntary or involuntary, it can take the form of games... It can mobilize very different types of motivations. In libraries, crowdsourcing is not without consequences on the economic model. It deploys many criticisms and is sometimes qualified as uberization of libraries.

Keywords : digital libraries, digitalisation, crowdsourcing, crowdfunding, ludification

Cette présentation générale au sujet du *crowdsourcing* dans les bibliothèques résulte d'un travail de thèse qui a été soutenue récemment [1], publiée sous la forme d'un livre [2] mais aussi résumée sous la forme d'un état de l'art [3].

Cette thèse a été rédigée sous Google Doc et rendue accessible en temps réel *via* le site Web bibliotheque-numerique.fr à la direction de la thèse comme à l'ensemble des personnes intéressées par ce sujet et qui pouvaient également y ajouter, de manière participative, des suggestions de corrections, de lectures ou des commentaires. Les contributeurs ont été remerciés sur le site Web, en fonction de leurs contributions, mais également dans la partie « Remerciements » de la thèse. À la question récurrente du

1. Andro, Mathieu (2016) *Crowdsourcing et bibliothèques numériques : expérimentations autour de Numalire, projet de numérisation à la demande par crowdfunding*, Thèse de doctorat en sciences de l'information et de la communication, Paris 8. Accessible sur `https://hal.archives-ouvertes.fr/tel-01384692/document`, consulté le 25 juillet 2020.

2. *Id.* (2017) *Bibliothèques numériques et crowdsourcing*, Paris, ISTE.

3. *Id.* ; Saleh, Imad (2017) « Bibliothèques numériques et crowdsourcing : un état de l'art », *Intelligence collective et archives numériques : Vers des écosystèmes de connaissances*, S. Szoniecky et N. Bouhaï, Paris, ISTE. p. 135-160

risque d'être plagié, nous avons estimé que les opportunités de collaborations étaient plus importantes que les risques. En effet, l'historique très détaillé du Google Doc aurait pu nous permettre de revendiquer assez facilement l'antériorité et l'originalité de certaines de nos analyses.

De manière introductive, et afin d'interpeller notre auditoire au sujet du *crowdsourcing*, nous souhaiterions reprendre la comparaison de Clay Shirky [4]. Cet auteur estime, en effet, que l'encyclopédie participative Wikipedia a nécessité environ 100 millions d'heures de travail bénévole afin d'être édifiée. Mais, dans le même temps, si on considère le temps libre que les citoyens américains consacrent à regarder la télévision, à savoir 200 billions d'heures par an, on peut en conclure que ces citoyens américains pourraient édifier 2000 projets comme Wikipedia chaque année s'ils consacraient leur temps à ce type de bénévolat plutôt qu'à regarder la télévision. Le *crowdsourcing* est donc une technologie dont le potentiel peut laisser rêveur.

1 Les origines historiques du concept

Il est possible de faire remonter l'origine historique du *crowdsourcing* aux premiers concours et appels à participation lancés par les États afin de résoudre des problèmes. Ainsi, le gouvernement anglais proposa, *via* le Longitude Act de 1714, 20 000 livres de récompense à qui parviendrait à déterminer la longitude d'un navire suite à l'accident de Cloudesley Shovell aux îles Scilly en 1707. Parmi la multitude de concurrents se trouvaient des scientifiques renommés et établis comme Cassini ou Newton, mais ce fut John Harrisson, un horloger, qui remporta le concours [5]. À cette époque, comme aujourd'hui, la puissance publique, consciente de ses limites, a fait appel à la puissance extérieure auprès de foules d'intelligences.

L'invention du libre-service par Aristide Boucicaut et le magasin *Au bon marché* au milieu du XIX^{e} siècle pourrait également être considérée comme un modèle précurseur du *crowdsourcing*. En effet, le travail du commerçant a été ainsi externalisé auprès du consommateur qui a eu accès directement aux marchandises sans passer par l'intermédiaire d'un magasinier. Dans le domaine des bibliothèques, ce modèle a trouvé son application dans le libre accès à la bibliothèque Sainte-Geneviève où l'architecte Labrouste conçoit qu'une partie du fonds sera accessible ainsi sans médiation. Un peu plus tard, en 1895, c'est à la bibliothèque publique de Clerkenwell que James Duff Brown conceptualise et expérimente le modèle du libre accès. Une partie du travail du bibliothécaire et du magasinier est ainsi externalisée auprès du lecteur.

Cette intégration du consommateur dans le procès de production se poursuit historiquement avec le modèle du « juste à temps » proposé dans les années 1950 par Taiichi Ōno au sein de l'entreprise Toyota. Il s'agit de produire en flux tendu, sans stocks ni invendus, directement en fonction de la demande. Et, ce faisant, il s'agit d'externaliser auprès du consommateur la décision de mise en production elle-même.

4. Shirky, Clay (2008) *Here comes everybody : the power of organizing without organizations*, London, Penguin Books.

5. Lakhani, Karim R. (2013) « Using the crowd as an innovation partner », *Harvard Business review*.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

Ce modèle du « juste à temps » se retrouve aujourd'hui au cœur de la numérisation et dans l'impression à la demande (*print on demand*) de livres [6].

La souscription publique – très populaire au XIX[e] siècle afin de financer de manière participative des monuments comme la statue de la Liberté qui avait mobilisé 125 000 donateurs en 1884, ou encore la publication de livres – est elle-même également l'ancêtre du *crowdfunding*, une forme de *crowdsourcing* qui fait appel non pas aux ressources intellectuelles ou à la force de travail des foules, mais à leurs ressources financières.

Avec ces multiples exemples, on peut constater que le *crowdsourcing* tire ses origines de modèles économiques relativement anciens.

2 Définition et taxonomie

Aujourd'hui, le *crowdsourcing* pourrait être défini comme l'externalisation (*outsourcing*) du travail, des compétences, des connaissances ou du financement (*crowdfunding*) auprès de foules (*crowd*) d'internautes.

De notre point de vue, on peut distinguer plusieurs formes de *crowdsourcing*. Il y a, tout d'abord, le *crowdsourcing* explicite classique qui consiste à faire simplement appel à du bénévolat volontaire et gratuit sur le Web.

Ensuite, le *crowdsourcing* rémunéré consiste à recruter des foules d'internautes qui sont également tout à fait volontaires pour travailler, qui le font non pas bénévolement, mais en échange d'une rémunération. Pour cette forme, on peut évoquer la plate-forme Amazon Mechanical Turk Marketplace développée dès 2005. Ce nom est directement inspiré d'un automate joueur d'échecs du XVIII[e] siècle, réputé doté d'une intelligence artificielle alors qu'en réalité une intelligence humaine et bien naturelle était cachée à l'intérieur. Le *crowdsourcing* permet en effet de mobiliser l'intelligence humaine au profit de projets irréalisables jusqu'à présent et au bénéfice d'objectifs encore irréalisables par l'intelligence artificielle. L'Amazon Mechanical Turk Marketplace rassemble un demi-million de travailleurs dans près de 200 pays dans le monde. En avril 2010, par exemple, ce sont 6 701 406 microtâches qui ont été proposées sur cette plate-forme par 9 436 sociétés et pour une valeur marchande de 529 259 $. Cette forme de *crowdsourcing* permet, en effet, de recruter en quelques clics des armées de travailleurs de compétences variées et qui, en raison de leur nombre, risquent moins le *burn out*. Ces travailleurs jouissent de la liberté de travailler où ils veulent, quand ils veulent, pour qui ils veulent et autant qu'ils le veulent. Mais tout cela dans la limite de rétributions dont le niveau n'est régulé que par la loi de l'offre et de la demande et qui peuvent être très basses. Ce modèle est donc parfois critiqué comme de l'exploitation et comme une « ubérisation » du travail.

Une autre forme de *crowdsourcing* est celle que nous qualifierons, à la suite de Harris [7], de *crowdsourcing* implicite. Il s'agit d'un travail inconscient et involontaire produit

6. Andro, M., Klopp, Sophie (2015) « L'impression à la demande et les bibliothèques », *Bulletin des Bibliothèques de France,* contribution du 13 février 2015.

7. Harris, Christopher Glenn (2013) *Applying human computation methods to information science*, PhD dissertation, University of Iowa.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

par les foules d'internautes grâce aux données et aux traces qu'ils laissent sur le Web. L'un des exemples le plus remarquables d'une application de cette forme de *crowdsourcing* au bénéfice de la numérisation des bibliothèques est reCAPTCHA [8], cette interface qui vous demande de saisir deux mots déformés afin de pouvoir créer un compte sur un site Web en prouvant que vous êtes bien un humain et pas un robot. L'un des deux mots ressaisis a bien cette fonction. En ressaisissant l'autre mot, vous travaillez en réalité, et peut-être sans le savoir, à corriger l'OCR de Google Books ou de Google Maps. En effet, Google a numérisé plus de 30 millions de livres. Ces livres ont fait l'objet de traitement de reconnaissance de caractères (OCR) afin de convertir les images de textes en textes indexables, cherchables et dans lesquels des extractions de données peuvent être réalisées (*text mining*). Malheureusement, selon la qualité du texte d'origine ou de la numérisation, ces traitements introduisent des erreurs. Les mots suspects, c'est-à-dire qui n'ont pas d'occurrence dans des dictionnaires de mots, sont ainsi envoyés aux internautes pour être corrigés, grâce au reCAPTCHA, en confrontant les saisies des internautes entre elles. Ainsi, on considérera que si trois internautes dans le monde ressaisissent le même mot exactement de la même manière, cela signifie que ce mot est désormais corrigé. Dans le cas contraire, il faudra faire appel à des internautes supplémentaires. Chaque année, ce sont environ 200 millions de mots dont 88 validés qui sont ainsi corrigés par les internautes, soit 12 000 heures et 1147 ebooks corrigés à un taux de qualité de 99,1 %. Si cette OCR brute était corrigée dans un pays en voie de développement comme le font les bibliothèques françaises qui corrigent de l'OCR, cela reviendrait à près de 46 millions d'euros par an d'après nos calculs [9]. Il s'agit d'une valeur importante produite inconsciemment par les internautes. Néanmoins, grâce à ce qui ressemble fort à une externalité positive, la qualité de Google Books qui est accessible gratuitement s'en trouve ainsi améliorée au bénéfice de tous. Et on peut considérer qu'il était assez économique et écologique d'utiliser ainsi le travail nécessaire à une vérification de sécurité au bénéfice de la plus grande bibliothèque numérique au monde. Toutefois, il faut néanmoins signaler que reCAPTCHA est désormais un système dépassé et de moins en moins fiable sur le plan de la sécurité.

Il existe aussi une forme intermédiaire entre le *crowdsourcing* volontaire et involontaire. Il s'agit de la « gamification » sous la forme de GWAP (*Games With A Purpose*). La gamification consiste, en effet, à faire travailler les internautes sous la forme de jeux [10]. C'est une forme intermédiaire, car, si l'internaute est conscient de contribuer, sa motivation n'est pas strictement celle d'un bénévole puisqu'il s'agit aussi de jouer. Dans le domaine des bibliothèques, on peut citer le jeu Digitalkoot de 2011 qui consistait à demander à des internautes de saisir des mots à partir des images numérisées de ces mots. À chaque mot correctement saisi, par confrontation avec les saisies d'autres joueurs, une brique nouvelle d'un pont venait s'ajouter aux précédentes afin de permettre à de petites taupes de traverser une rivière. À chaque erreur, une brique du

8. Von Ahn, Luis ; Maurer, Benjamin ; Mcmillen, Colin ; Abraham, David ; Blum, Manuel (2008) « reCAPTCHA : Human-Based Character Recognition *via* Web Security Measures », *Science,* n° 321, p. 1465-1468.

9. Andro, M. et Saleh, I. (2015) « La correction participative de l'OCR par *crowdsourcing* au profit des bibliothèques numériques », *Bulletin des Bibliothèques de France,* contribution du 16 juin 2015.

10. *Eid.* (2015) « Bibliothèques numériques et gamification : panorama et état de l'art », *I2D – Information, données & documents* vol. 52, n° 4, p. 70-79.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

pont explosait, engendrant de malheureuses noyades. Comme dans un jeu d'arcade classique, les décors et les difficultés augmentaient à chaque niveau. D'autres jeux dérivés étaient également proposés sur ce même modèle par ce projet, mais aussi par d'autres.

Enfin, une dernière forme de *crowdsourcing* nous semble être le *crowdfunding*. En effet, avec cette dernière forme, il ne s'agit pas de demander des ressources « travail » aux foules d'internautes mais de faire plutôt appel à leurs ressources financières. Dans le domaine des bibliothèques, nous avons ainsi expérimenté un projet de numérisation à la demande sous la forme de *crowdfunding* qui s'appelait Numalire [11]. Il s'agissait, pour les bibliothèques ne disposant pas d'un service de reprographie, de proposer aux internautes de financer la numérisation de livres pour leurs besoins ou sous la forme de mécénat. Ces livres étaient ensuite numérisés par un prestataire rémunéré par les internautes. Cela permettait aux bibliothèques d'offrir un service de reprographie de qualité professionnelle sans avoir à en supporter le coût, et à leurs usagers de bénéficier d'un service de reprographie généralement inexistant. Cela permettait aussi de partager la politique d'acquisition des bibliothèques numériques avec les internautes et de constituer des bibliothèques numériques qui soient l'œuvre des internautes eux-mêmes, c'est-à-dire directement centrées sur leurs besoins. Cela permettait également, pour les bibliothèques, d'externaliser le travail très ingrat consistant à identifier les documents méritant d'être numérisés et à vérifier que les documents n'avaient pas déjà été numérisés. Enfin, le projet aurait également pu offrir à des sponsors et à des investisseurs la possibilité de financer la numérisation de livres dans l'espoir de connaître un retour sur investissement sous la forme de trafic Web à la manière de publicités Google Adwords, mais dont la durée n'aurait pas été limitée. Ainsi, au-delà d'un certain nombre de vues d'une mention, d'un logo et de clics sur les métadonnées d'un document numérisé, le sponsor aurait rentabilisé le coût de la numérisation. De cette manière, l'argent public aurait pu être dédié à la numérisation de livres à haute valeur scientifique, historique ou de conservation, et l'argent privé aurait financé la numérisation de livres susceptibles de générer du trafic Web ou de répondre plus directement aux besoins d'internautes.

Nous venons de dresser ainsi une taxonomie du *crowdsourcing* avec le *crowdsourcing* explicite qui a recours au travail volontaire des internautes, et le *crowdsourcing* implicite qui a recours au travail involontaire des internautes (comme reCAPTCHA). Dans le *crowdsourcing* explicite, on peut distinguer le *crowdsourcing* explicite gratuit, c'est-à-dire le recours au travail volontaire des internautes bénévoles (comme TROVE, Transcribe Bentham, Flickr, The Commons) et le *crowdsourcing* explicite rémunéré, c'est-à-dire le recours au travail volontaire des internautes rémunérés (comme l'Amazon Mechanical Turk Marketplace ou FouleFactory en France). Entre le *crowdsourcing* implicite et le *crowdsourcing* explicite, il existe une forme intermédiaire avec la gamification (ou « *human computation* » ou encore « *games with a purpose* »), c'est-à-dire le recours au travail des internautes sous la forme de jeux (comme Digitalkoot ou Art Collector). Et pour finir, nous avons évoqué le *crowdfunding*, c'est-à-dire le re-

11. Andro, M. ; Rivière, Pauline ; Dupuy-Olivier, Anaïs ; Gropallo, Filippo ; Maingreaud, Denis (2014) « Numalire, une expérimentation de numérisation à la demande du patrimoine conservé par les bibliothèques sous la forme de financements participatifs (*crowdfunding*) », *Bulletin des Bibliothèques de France,* contribution du 2 octobre 2014.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

cours aux contributions financières des internautes (comme Ebooks on Demand ou Numalire).

Cette taxonomie pourrait faire l'objet de discussions. Ainsi, le *crowdsourcing* implicite pourrait être considéré comme n'étant pas du *crowdsourcing* en raison de son caractère involontaire. Restreindre le *crowdsourcing* au seul bénévolat en ligne nous semblerait une erreur à l'heure où ces projets se multiplient : il est possible que le *crowdsourcing* implicite et la gamification représentent justement l'avenir du *crowdsourcing*. Par ailleurs, il nous semblerait assez difficile de ne pas qualifier le projet reCAPTCHA de projet de *crowdsourcing*. D'autres pourraient également souhaiter exclure le *crowdsourcing* rémunéré pour des raisons idéologiques, mais la frontière entre rémunération symbolique, en nature ou en espèces reste délicate et il n'est guère évident, sur le plan de l'éthique, de préférer le travail gratuit au travail rémunéré.

3 Les clés du succès

Maintenant que nous avons abordé les origines, la définition et les différentes formes de *crowdsourcing*, nous allons évoquer les questions importantes pour tout projet de *crowdsourcing*.

Pour pouvoir recruter, les projets vont devoir investir dans des campagnes de communication dans les médias, les réseaux sociaux, à travers l'organisation d'évènements... Pour pouvoir maintenir les contributeurs, ils vont devoir investir dans du *community management*. Pour pouvoir réellement exploiter les données, il va falloir former les contributeurs, contrôler la qualité des données qu'ils produisent. La réintégration des données produites doit effectivement être imaginée très en amont. Enfin, les projets seront évalués par des tableaux de bord de statistiques et d'indicateurs, il faudra comparer les coûts et les bénéfices obtenus avec les coûts et les bénéfices qui auraient été ceux d'une externalisation auprès d'un prestataire plus traditionnel.

La question des motivations des internautes est une question centrale. On distingue les motivations intrinsèques qui poussent l'individu à réaliser une activité pour la satisfaction que cette activité lui procure en soi, et les motivations extrinsèques qui poussent l'individu à réaliser une activité pour obtenir un résultat extérieur à l'activité qui n'est qu'un moyen de l'obtenir. Parmi les motivations intrinsèques, on rencontre des motivations individuelles comme l'intérêt pour un sujet, la recherche de développement personnel autour d'activités diversifiées et réalisées de manière autonome, la distraction, l'amusement, l'innovation, la résolution de problèmes, le fait de se prouver quelque chose ou d'améliorer son estime de soi, l'esprit de compétition ou encore l'impression d'avoir du pouvoir sur les choses. Parmi les motivations intrinsèques, on rencontre aussi des motivations plus collectives avec le sentiment d'appartenance à une communauté, le fait de rencontrer d'autres contributeurs, d'agir par altruisme ou encore par sentiment de redevabilité. Parmi les motivations extrinsèques cette fois, on rencontre la rétribution immédiate sous la forme de rémunérations symboliques (classements, remerciements publics...), ou plus réelles de cadeaux (livres, abonnements, bons d'achat, formations...) ou d'avantages (visites personnalisées de la bibliothèque), mais aussi la rétribution future sous la forme d'une amélioration de son CV pour trouver un emploi, de sa e-reputation ou encore la recherche d'une réciprocité.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

En fonction du type de *crowdsourcing*, on utilisera plutôt tel ou tel ressort pour récolter des contributions des internautes. On observe parfois qu'en faisant appel à un type de motivation au détriment d'autres, on peut avoir un effet contraire à celui recherché. Ainsi, dans un domaine totalement étranger aux bibliothèques, on constate parfois, si on rémunère les donneurs de sang, une diminution des quantités données ; ou, si on rémunère trop les *crowdsourcers* de l'*Amazon Mechanical Turk Marketplace*, une diminution de la qualité. Ainsi, si on récompense des enfants pour leurs dessins, on peut constater une augmentation de leur production, mais au détriment de la qualité de leurs œuvres et, surtout, au détriment du plaisir qu'ils éprouvent à dessiner.

On constate aussi, bien souvent, que pour les projets culturels, il est préférable de recruter quelques bénévoles bien identifiés et bien motivés plutôt que des foules sans visages. En effet, on observe pour tous les projets qu'une minorité d'internautes produit la grande majorité des contributions. Aussi, on devrait plutôt parler dans ce cas de *communitysourcing* ou de *nichesourcing*[12].

4 Les critiques

La plupart des opposants au *crowdsourcing* mettent en avant la question de la malveillance et de la qualité. Le lambda se sentirait ainsi autorisé à s'approprier individuellement un bien commun et un patrimoine collectif en ajoutant des commentaires profanes et des tags triviaux. Pourtant, des dispositifs, de confrontation de données notamment, permettent d'obtenir des données de qualité. Par ailleurs, s'agissant par exemple de l'indexation (folksonomie) d'un document numérisé, il est fort probable qu'un internaute qui consulte le document connaisse mieux le sujet et soit donc plus efficace pour qualifier le document avec des mots clés libres qu'un bibliothécaire – malgré son excellente culture générale et son vocabulaire contrôlé.

Un autre argument généralement mis en avant pour plaider contre le *crowdsourcing* est d'ordre plus politique, social et idéologique. Le *crowdsourcing* serait le stade suprême de l'externalisation. Il viserait à remplacer les salariés par des internautes bénévoles ou des internautes sous-payés (« servuction »). Il exercerait aussi une concurrence déloyale sur les prestataires traditionnels et pourrait générer un désengagement des travailleurs comme des employeurs, et même des escroqueries. Mais à y regarder de plus près, rien n'interdit aux bibliothèques de rémunérer les contributeurs sur le Web et de bien les rémunérer.

En France, les bibliothèques fondent leur modèle sur un modèle aristocratique d'experts ; elles ont été, historiquement, davantage tournées vers la conservation de collections, parfois issues de confiscations révolutionnaires, que vers la satisfaction des besoins des lecteurs. Dans ces conditions, le *crowdsourcing* peut être considéré comme une remise en cause et comme une dévalorisation de l'expertise des bibliothécaires à qualifier les documents, à indexer l'information et même à sélectionner les documents à acquérir. Ainsi, traditionnellement, une bibliothèque fonctionnait selon un modèle d'acquisition de documents en anticipant les besoins de ses lecteurs (« *just in case* »). La numérisation à la demande par *crowdfunding* est un modèle radicalement contraire

12. Causer, Tim ; Wallace, Valerie (2012) « Building A Volunteer Community : Results and Findings from Transcribe Bentham », *Digital Humanities quarterly,* vol. 6, n° 2.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

puisque la bibliothèque numérique est constituée directement par le lecteur (« *just in time*[13] »).

Nous pourrions également critiquer certains projets de *crowdsourcing* sur le plan très rationnel de leur rentabilité. Si on obtient de moins bons résultats avec un coût supérieur à celui d'un prestataire, c'est peut-être tout simplement que le *crowdsourcing* n'est pas la meilleure solution. Et on constate effectivement que très rares sont les projets de *crowdsourcing* qui communiquent autour des coûts. On constate aussi que de nombreux projets ne recherchent, à travers le *crowdsourcing*, qu'à flatter leur communication institutionnelle autour d'un sujet à la mode. Ainsi, d'après une étude menée par l'OCLC sur les projets de *crowdsourcing*[14], il semble que ceux qui utilisent réellement les données produites par les internautes sont malheureusement loin d'être majoritaires. Il s'agit donc dès lors d'une simple démocratisation de la culture ou d'un « *empowerment* », quand il ne s'agit pas de marketing autour de l'activité de l'institution ou de celle d'un de ses agents.

On remarque enfin que certaines institutions choisissent de consacrer leurs ressources à développer des plates-formes et à communiquer pour recruter des internautes dont le travail sera ensuite totalement gratuit, quand d'autres institutions consacrent leurs ressources à rémunérer les internautes sur des plates-formes existantes et déjà fréquentées par les internautes. Dans le premier cas, l'intérêt du producteur semble être privilégié. Dans le second cas, les intérêts des contributeurs et des lecteurs semblent prévaloir.

5 Conclusion

Nous sommes bien conscients que certains des propos qui précèdent, assez marginaux, sont susceptibles de provoquer le débat. Nous constatons, en effet, que le *crowdsourcing* suscite des polémiques et que nombre d'auteurs, notamment en France, prennent ainsi le parti de ne pas parler de certaines formes de *crowdsourcing* et de certains sujets pour des raisons idéologiques[15]. Nous avons toutefois fait le choix de n'écarter aucune forme de *crowdsourcing* ni aucun point de vue, laissant chacun libre de se faire son opinion en disposant du maximum d'éléments pour cela.

Au cours de cet exposé d'introduction au *crowdsourcing* en bibliothèque, nous avons évoqué les origines historiques du *crowdsourcing*, sa définition, ses différentes formes, les clés du succès des projets et les critiques qui en sont faites.

Le *crowdsourcing* est bien, comme son nom l'indique, une externalisation du travail auprès des lecteurs des bibliothèques, qui engendre une disparition des frontières entre amateur et expert, entre bénévole et professionnel, entre celui qui lit et celui qui écrit, et plus généralement entre celui qui consomme et celui qui produit. Avec le

13. Klopp, Sophie (2014,) *Numérisation et impression à la demande en bibliothèque : un panorama*, Mémoire de conservateur, Villeurbanne, ENSSIB.

14. Smith-Yoshimura, Karen ; Godby, Carol Jean ; Hoffler, Helice, *et al.* (2011) *Social Metadata for Libraries, Archives, and Museums : Survey analysis*, Dublin, OCLC Research.

15. *Id.* ; Saleh, Imad (2015) « Bibliothèques numériques et *crowdsourcing* : une synthèse de la littérature académique et professionnelle internationale sur le sujet », dans Zreik, Khaldoun ; Azemard, Ghislaine ; Chaudiron, Stéphane ; Darquié, Gaétan (dir.), *Livre post-numérique : historique, mutations et perspectives*, colloque international sur le document électronique (CiDE.17), Paris, Europia.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

crowdsourcing, l'usager devient à la fois consommateur d'un service et, en même temps, producteur d'une bibliothèque numérique.

Il s'agit peut-être aussi tout simplement d'une « ubérisation » des bibliothèques, c'est-à-dire d'une remise en cause de la médiation et du monopole de professions réglementées. Mais pour terminer sur une note positive, nous ne pouvons qu'espérer que le *crowdsourcing* nous permette de rendre possibles des projets jusque-là inimaginables (Holley, 2010) et que la collaboration avec des amateurs apporte des ruptures innovantes dans nos bibliothèques. En effet, les non-professionnels ne cherchent pas à reproduire les modèles établis du métier et sont parfois susceptibles d'apporter d'heureuses coïncidences d'« innovations par les usagers » (*user innovation*) [16].

16. Von Hippel, Eric (2005) *Democratizing innovation*, Cambridge, MIT Press.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850