

# Le *crowdsourcing* ou la participation sociale et scientifique d'un grand nombre d'individus à l'interprétation des données de recherche en sciences de l'éducation

Thierry **Karsenti**, Ph.D.   
@ThierryUdM  
Directeur du CRIFPE  
Chaire de recherche du Canada  
sur les technologies en éducation

doi:10.18162/fp.2015.a65

## CHRONIQUE • Technologies en éducation

### Qu'est-ce que le *crowdsourcing*?

À une époque où les réseaux sociaux sont omniprésents dans la plupart des sphères de la société, il est de plus en plus fréquent d'entendre le terme *crowdsourcing* (de l'anglais, approvisionnement par la foule ou par un grand nombre de personnes). Le terme *crowdsourcing* a été avancé pour la première fois en 2006 par Howe qui l'a défini comme étant « *an online, distributed problem-solving and production model [...]* » (p. 1). Cette stratégie sociale de recherche vise à trouver une réponse à un défi rencontré, par l'aide directe ou indirecte d'un grand nombre d'individus contactés par l'entremise des réseaux sociaux. Howe avance que les progrès technologiques, notamment du Web 2.0 et des réseaux sociaux, permettent à des entreprises de profiter des compétences de la population, d'où le mot-valise ou néologisme *crowdsourcing*. Il s'agit possiblement de l'un des domaines les plus émergents de la gestion des connaissances. Le terme s'est d'ailleurs rapidement répandu auprès de centaines d'entreprises, d'auteurs et de médias, comme *raccourci* pour désigner une nouvelle tendance à utiliser la collaboration de masse permise par les réseaux sociaux, pour atteindre des objectifs économiques, culturels, sociaux ou scientifiques. Avec le *crowdsourcing*, l'appel peut-être ciblé – si l'on souhaite par exemple faire appel à certaines personnes en particulier – ou encore ouvert, si c'est un plus large public qui est visé par la demande. La participation peut être collaborative si un tout est construit à la fin du processus, un peu à la manière dont Wikipédia est conçu. La participation peut aussi s'effectuer en parallèle, dans le cas où les réponses reçues sont comparées et analysées pour n'en retenir que les meilleures par exemple. C'est ce qui a amené Van Ess (2010)

à présenter le *crowdsourcing* comme méthode servant à canaliser les besoins d'experts pour résoudre un problème. Le *crowdsourcing* se retrouve dans divers domaines. Par exemple, le projet Stardust@home<sup>1</sup> permet à la NASA d'avoir l'aide du grand public pour chercher des impacts de poussière spatiale. Le *crowdsourcing* est également répandu dans les systèmes d'information géographiques, avec de nouveaux outils collaboratifs comme OpenStreetMap<sup>2</sup>. On parlera pour ces deux exemples respectivement de *science citoyenne* ou de *cartographie citoyenne*. Le financement participatif (en anglais *crowdfunding*) est aussi une expression décrivant une stratégie financière qui fait appel à un grand nombre de personnes pour financer un projet, par exemple l'initiative Dev4X<sup>3</sup>.

## Le *crowdsourcing* en sciences sociales

Néanmoins, à notre connaissance, la pratique technosociale de faire appel à un grand nombre d'individus pour participer à l'interprétation des données de recherche ne semble pas encore répandue en sciences humaines. La stratégie du *crowdsourcing* qui pourrait permettre une meilleure interprétation des données de recherche recueillies, à partir d'hypothèses formulées par un grand nombre d'individus, gagnerait peut-être à être explorée.

En sciences humaines, on retrouve certes quelques parties de recherche qui font appel à divers individus ou stratégies. On retrouve par exemple depuis bon nombre d'années la pratique de la prépublication. Elle peut être définie comme un genre de brouillon d'un article scientifique qui n'a pas encore fait l'objet d'une publication dans une revue. Cette pratique est d'abord née à cause des importants délais inhérents aux publications scientifiques, imputés notamment au processus d'évaluation par les pairs. Cette pratique a aussi connu un essor important pour permettre à certains chercheurs de partager rapidement des résultats de recherche avec la communauté scientifique, et ce, afin de recevoir des commentaires (d'autres chercheurs), ce qui peut s'avérer fort utile en vue d'accroître les chances d'un texte d'être accepté dans une revue scientifique. Avec l'arrivée d'Internet et des réseaux sociaux, les prépublications ont connu un essor important, et on retrouve même certaines bases de données qui se spécialisent dans leur diffusion, par exemple arXiv.org. Ces bases de données sont toutefois bien plus populaires en sciences appliquées comme la physique qu'elles ne le sont en sciences humaines. Même *Nature* a conçu son propre répertoire de manuscrits en prépublication (*Nature Precedings*<sup>4</sup>), avec le même objectif que la base arXiv.org. Néanmoins, la pratique des prépublications vient également avec son lot de désavantages, notamment en ce qui a trait à l'absence de reconnaissance des pairs pour la carrière professorale.

Le *crowdsourcing* diffère énormément de la pratique du *preprint* retrouvée dans le monde scientifique. Il ne s'agit pas de publier une version préliminaire d'un manuscrit pour recevoir l'avis de collègues-chercheurs, mais plutôt de faire appel aux réseaux sociaux pour aider le chercheur à interpréter des données de recherche.

On retrouve également une autre forme de validation en sciences humaines : la triangulation dans l'interprétation des données de recherche (voir Karsenti et Savoie-Zajc, 2011). La triangulation peut être vue comme une stratégie de recherche au cours de laquelle le chercheur superpose ou combine plusieurs perspectives, qu'elles soient de théories, de méthodes, voire de participants à la recherche (voir Karsenti et Savoie-Zajc, 2011). Le processus de triangulation diffère toutefois du *crowdsourcing* parce

qu'il possède des origines positivistes. En effet, dès 1959, Campbell et Fiske proposaient de réaliser des recherches en introduisant plusieurs instruments de mesure pour compenser les limites de chacun : « *The achievement of useful hypothetically realistic constructs in a science requires multiple methods focused on the diagnosis of the same construct from independent points of observation through a kind of triangulation* » (p. 81). La notion a été reprise par Denzin en 1970 et appliquée à la recherche qualitative : « *The rationale for this strategy is that the flaws of one method are often the strengths of another, and by combining methods; observers can achieve the best of each, while overcoming their unique deficiencies* » (p. 308).

La triangulation dans la recherche en sciences humaines poursuit en fait deux objectifs. Le premier est de permettre au chercheur d'explorer le plus de facettes possible du problème étudié en recueillant des données qui vont faire ressortir des perspectives diverses. Ceci permettra de dégager une compréhension riche du phénomène analysé. Le second vise à mettre la triangulation au cœur du processus de coconstruction des connaissances et à supporter l'objectivation du sens produit pendant la recherche. Habermas (1987) nomme ce processus le « procès d'intercompréhension ». Il fait référence à la dynamique qui s'établit entre des personnes qui souhaitent communiquer et qui, au cours de leur contact, cherchent à être comprises l'une de l'autre. Pareille compréhension entre les interlocuteurs prend pour mesure la vérité, la justesse et la véracité entre le langage utilisé et les mondes des acteurs, construits par leur expérience, leurs savoirs, leur culture, leur biais. On peut ainsi distinguer entre plusieurs types de triangulation (voir Karsenti et Savoie-Zajc, 2011) qui vont se rattacher à l'un ou à l'autre de ces objectifs (triangulation théorique, triangulation des méthodes, triangulation du chercheur, triangulation des sources, triangulation par l'analyse, etc.). Relativement au deuxième objectif de la triangulation, celui d'objectiver la démarche de coconstruction des savoirs, la triangulation indéfinie supporte ce souci de retour aux participants à la recherche et suscite la discussion autour des constructions de sens émergentes. Ce retour s'effectue surtout pendant l'analyse des données, alors que le chercheur s'efforce de susciter les réactions des participants à propos des constructions qui s'élaborent autour de leurs propos recueillis.

Le crowdsourcing dans la recherche en sciences sociales semble plus s'apparenter à cette forme de triangulation. Sauf que le chercheur ne se limite alors pas à susciter les réactions des participants à sa recherche (pour valider, par exemple, la synthèse d'une entrevue réalisée), mais bien d'un grand nombre d'individus, notamment par l'entremise des technologies et des réseaux sociaux. Cette pratique peut d'ailleurs être fort utile dans un contexte où le chercheur rencontre de grandes difficultés à interpréter les données recueillies, ce qui a justement été notre cas dans l'exemple présenté ci-après. En faisant appel à un grand nombre d'individus, le chercheur accepte alors de recevoir à la fois des commentaires de pairs (forme de validation scientifique), mais aussi des commentaires d'autres individus qui ne sont pas nécessairement des experts dans le domaine, mais qui ont néanmoins une opinion à partager (forme de validation sociale). Dans notre cas, il a été plus particulièrement question d'une validation par des praticiens (des enseignants). L'usage du *crowdsourcing* dans la recherche en sciences sociales peut donc s'apparenter à une forme de triangulation à la fois sociale et scientifique. Cette pratique pourrait s'avérer à la fois novatrice et fort intéressante pour le chercheur en sciences humaines.

## Usage du *crowdsourcing* dans l'interprétation de données de recherche en sciences de l'éducation

Le *crowdsourcing* nous est apparu utile lors d'une récente étude sur l'usage de l'iPad pour l'écriture de textes de quelque 300 mots. Lors de cette étude, réalisée auprès de 143 élèves âgés de 12 à 14 ans (enquête par questionnaire, observations de classe vidéographiées, entrevues individuelles et de groupes, analyse des traces informatiques, corrections des épreuves de français écrit, etc.), nous éprouvions énormément de difficultés à trouver des pistes d'interprétation aux données recueillies. Il nous est donc apparu intéressant d'envoyer des extraits de plusieurs de ces données, de façon anonyme, à un très grand nombre d'individus, notamment par l'entremise des réseaux sociaux (surtout via Twitter et les quelque 4600 abonnés au compte @ThierryUdM). Pourquoi avoir fait appel au *crowdsourcing*? Parce que les données recueillies étaient surprenantes. Nous avons d'abord été complètement étonnés de constater à quel point l'usage d'un clavier externe était superflu pour 98,6 % des jeunes interrogés. Les résultats préliminaires de cette même enquête ont également participé à nourrir un questionnement sur l'apprentissage ou non du doigté sur le clavier. Les adultes (enseignants, formateurs et autres praticiens) interrogés étaient tous en faveur de cet apprentissage, puisque plusieurs étaient, eux-mêmes « *passés par là* » (extrait d'entrevue avec un enseignant). Et la plupart étaient aussi, pour les mêmes raisons, en faveur de l'usage d'un clavier lors de longues rédactions écrites. Mais pourquoi? Pour écrire plus vite? Nos résultats ont montré clairement que la vitesse était plutôt un des points forts des jeunes observés. Il faut d'ailleurs regarder certaines de ces délicieuses vidéos anonymisées<sup>5</sup> pour voir à quel point la vitesse est un atout pour eux. Plusieurs dépassaient d'ailleurs les 80 mots/minute avec uniquement « deux pouces », voire parfois avec un pouce et un index. Et on parle de rédactions complètes, de plusieurs centaines de mots, réalisées en contexte scolaire, et écrites en quelques minutes à peine. Il s'agissait donc bien plus que de simples textos écrits avec deux pouces. Alors, pourquoi apprendre le doigté? Pour mieux écrire? Plusieurs de ces élèves qui écrivent à « deux pouces » arrivent à terminer une rédaction sans aucune faute de français (c'est ce qu'a révélé la correction des productions écrites). Pourquoi alors? Nous étions sans réponse. Nos sociétés occidentales ont déjà vécu deux grandes révolutions : le passage de l'oral à l'écrit, puis celui de l'écrit à l'imprimé. La troisième semblait résolument être le passage de l'imprimé aux nouvelles technologies, qui semble tout aussi majeure. Nous avons d'ailleurs trop peu de pistes qui nous apparaissaient intéressantes dans l'analyse de ces données. Et c'est là que la stratégie du *crowdsourcing* nous est apparue intéressante à explorer. La diffusion de plusieurs données de recherche, toutes anonymisées, entre le 15 juin et le 21 août 2015, a permis de recueillir quelque 2183 commentaires (la plupart reçus sur Twitter, mais aussi sur Facebook, par courriel, etc.). Quoique l'objectif de ce texte ne soit pas de présenter les résultats, il est intéressant de faire remarquer que 6 % des répondants étaient des experts (formateurs universitaires), 63 % des praticiens (enseignants, conseillers pédagogiques, administrateurs scolaires), et pour 31 % des répondants, il n'a pas été possible de déterminer leur statut. Cette méthode nous a donc permis de confronter les données recueillies auprès de 143 élèves à l'interprétation de quelque 2183 personnes, principalement des praticiens et des chercheurs. Nous avons analysé les réponses et hypothèses reçues à l'aide du logiciel *QDA Miner*, ce qui nous a même permis de catégoriser les solutions reçues. Mais cela fera l'objet d'un autre texte.

## Conclusion : le crowdsourcing en sciences sociales, une méthode à explorer?

Le *crowdsourcing* a semblé riche de potentialité, surtout parce que cette stratégie a permis de questionner à la fois un grand nombre de praticiens et d'experts sur l'interprétation de données recueillies dans un domaine encore nouveau : l'écriture sur tablette. Est-ce que cela a permis d'obtenir « la » bonne réponse, la « vraie » interprétation des données recueillies? Ce n'était pas l'objectif. Mais cela a définitivement participé à nous ouvrir à un vaste horizon d'hypothèses, d'explications, et de solutions potentielles. En outre, la *loi du nombre* offre une puissance d'analyse exceptionnelle qui permet, aussi, d'éliminer un grand nombre de réponses douteuses.

Et n'oublions pas non plus l'absence de coût inhérent au travail de ces quelque 2183 assistants de recherche du moment. Autrement dit, les participants volontaires ont offert, *a priori*, généreusement leurs compétences. Ils pouvaient estimer être en quelque sorte remboursés de leur travail par des résultats dont tout un chacun profitera, ou estimer que l'intérêt général du projet justifie leur participation. Mais il serait nécessaire d'être clair sur cette question, au moment de la sollicitation. Aussi, il aurait pu être nécessaire de s'assurer du statut exact de chacun des répondants. Mais, là encore, ce n'était pas l'objectif. La méthode du *crowdsourcing* comporte donc de nombreux avantages, et elle devra évidemment faire l'objet d'autres recherches. Plusieurs questions demeurent aussi : qu'en est-il, par exemple, de la propriété intellectuelle du travail bénévole des 2183 interprétateurs? Il aurait peut-être fallu que nous soyons plus précis dans l'appel. En d'autres termes, il faudra aussi, sans pour autant alourdir le processus, réfléchir à certaines règles éthiques pour éviter que cette pratique ne donne lieu à des dérives. Il ne faut pas non plus oublier que s'il s'agit de poursuivre un but scientifique, la validité et la légitimité d'un travail basé sur le *crowdsourcing* nécessitent un protocole scientifique et technique clair, transparent, rigoureux et crédible, surtout si l'on souhaite que cette pratique puisse intégrer les méthodologies de recherche en sciences humaines.

## Notes

- 1 <http://stardustathome.ssl.berkeley.edu/>
- 2 <https://www.openstreetmap.org>
- 3 <http://www.dev4x.com>
- 4 <http://precedings.nature.com>
- 5 <http://karsenti.ca/petitepoucette/>

## Références

- Campbell, D. T. et Fiske, D. W (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105. <http://dx.doi.org/10.1037/h0046016>
- Denzin, N. K. (1970). *The research act : A theoretical introduction to sociological methods*. Chicago, IL : Aldine.
- Howe, J. (2006). The rise of crowdourcing. *Wired*, (14), 6 juin. Repéré à [http://archive.wired.com/wired/archive/14.06/crowds.html?pg=1&topic=crowds&topic\\_set](http://archive.wired.com/wired/archive/14.06/crowds.html?pg=1&topic=crowds&topic_set)
- Karsenti, T. et Savoie-Zajc, L. (2011). *La recherche en éducation : étapes et approches* (3<sup>e</sup> édition). Montréal, QC : ERPI.
- Van Ess, H. (2010). *Harvesting Knowledge : how to crowdsource*. Communication présentée à ARD/ZDF Academy, Allemagne. Repéré à <http://fr.slideshare.net/searchbistro/harvesting-knowledge-how-to-crowdsource-in-2010>

## Pour citer cet article

- Karsenti, T. (2015). Le crowdsourcing ou la participation sociale et scientifique d'un grand nombre d'individus à l'interprétation des données de recherche en sciences de l'éducation. *Formation et profession*, 23(3), 112-116. <http://dx.doi.org/10.18162/fp.2015.a65>