

Open, Big, Collaboration: Trois utopies de l'innovation au XXIème siècle.

Fidelia Ibekwe-Sanjuan, Françoise Paquensegy

► To cite this version:

Fidelia Ibekwe-Sanjuan, Françoise Paquensegy. Open, Big, Collaboration: Trois utopies de l'innovation au XXIème siècle.. Evelyne Bredoux & Ghislaine Chartron. Colloque "Document numérique et Société". Open Data, Big Data: Quelles valeurs? Quels enjeux?, May 2015, Rabat, Maroc. De Boeck, pp.17-29, 2015, Information & Stratégie. <hal-01223138>

HAL Id: hal-01223138

<https://hal.archives-ouvertes.fr/hal-01223138>

Submitted on 2 Nov 2015

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Open, Big, Collaboration : trois utopies de l'innovation au 21^e siècle

Fidelia Ibekwe-SanJuan

IRSIC - EA 4262

Aix-Marseille Université

fidelia.ibekwe-sanjuan@univ-amu.fr

Françoise Paquienseguy

ELICO - EA 4147

Sciencespo Lyon

francoise.paquienseguy@sciencespo-lyon.fr

Introduction : vers un renouveau cybernétique ?

Trois idéaux ou utopies semblent stimuler l'innovation au XXI^e siècle : l'ouverture (*Open*) ; la participation et le gigantisme (*Big*). Si ces trois utopies rencontrent autant d'échos positifs auprès d'acteurs inscrits dans des logiques différentes, c'est que beaucoup pensent qu'elles ont la capacité d'instaurer un triangle vertueux et de devenir un levier libérant les énergies et accélérant le rythme des innovations. Pourtant, elles mêlent, en apparence, des idéologies contradictoires.

Breton et Proulx (2005) offrent une grille d'analyse des idéologies sous-jacentes au développement des technologies de l'information : elles puisent essentiellement dans la cybernétique de Norbert Wiener (1948) et s'organisent autour de trois courants : libertaire, libéral et régalien.

Au vu des effets dévastateurs des deux guerres mondiales, Wiener considérait que la fermeture et le secret étaient synonymes de sclérose des systèmes (organiques ou artificiels), qui s'en trouvaient figés. Il posait donc l'ouverture, la libre circulation de l'information et la libre communication comme des remparts contre le retour de la dictature et du chaos.

Le courant libertaire prône l'ouverture et la collaboration comme doctrines pour l'élaboration des biens et des connaissances. Au plan politique, l'idéal de l'ouverture est en toile de fond des mouvements citoyens et altermondialistes qui depuis les années 1980, militent pour la protection des biens communs et mondiaux, pour une société sans contrôle étatique où la circulation des biens et des personnes se ferait sans entraves et qui s'autorégulerait grâce aux nouvelles technologies. Il s'agit donc d'une forme de démocratie directe dont le *Parti Pirate*¹ est une illustration.

Soutenu par des mouvements politiques et citoyens, ce courant a permis l'essor d'une série de technologies de l'information dont l'ouverture des codes sources et des contenus devient la raison d'être, comme le témoigne l'inflation du préfixe « open » (Open Source, Open Archives, Open Journal, Open Edition, Open Education, Open Data, Open Innovation, Open Science ...).

Le deuxième courant (ultra)libéral s'oppose au premier dans la mesure où il valorise une marchandisation généralisée. Les technologies de l'information deviennent ici une opportunité pour investir dans des domaines jusque-là protégés des lois du marché tels la santé, l'éducation, les données personnelles, etc. Les offensives des géants du web, *via* l'emprise de leurs plateformes numériques, parviennent peu à peu à reprendre aux États de nombreuses prérogatives. Les innovations promises autour de l'Internet des objets comme les appareils de *self-tracking* ou de *quantified-self* collectent des données sur chaque aspect quantifiable de nos vies. Ces données ne manqueront pas d'alimenter des algorithmes utilisés par des sociétés commerciales pour calculer ensuite des grilles tarifaires d'assurance individuelle dans tous ces secteurs. A terme, c'est la notion même de social ou de solidarité qui risque d'être remise en question.

Dans certains secteurs, le modèle de l'ouverture est devenu la modalité principale de production des biens numériques. Ce modèle négocie désormais des formes de cohabitation avec l'idéologie libérale et marchande caractérisée par la fermeture et le brevetage.

Le troisième courant, régalien, représente les intérêts des États souhaitant contrôler l'évolution et les usages des technologies de l'information, d'abord vues comme un moyen d'assurer leur souveraineté et de sauvegarder leur suprématie. Il tente donc d'arbitrer entre les effets des courants libertaire et ultra-libéral, alors que ces deux courants sont "scellés dans une alliance"

¹ <http://partipirate.org>, consulté le 15 mai 2015.

militant pour l'absence totale de régulation étatique (Breton et Proulx, 2005). Les déclarations de certains dirigeants des industries du numérique accréditent cette thèse tant ils ne cachent pas leur immense mépris pour les politiques : "*In Silicon Valley, government is considered slow, staffed by mediocrities and ridden with obsolete rules and inefficiencies*"². Mis à part la Chine, la plupart des États semblent avoir perdu la bataille du contrôle des technologies de l'information au profit du courant ultra-libéral qui doit néanmoins composer avec le courant libertaire lui résistant par diverses formes, regroupées dans l'utopie de l'ouverture.

Conjuguées, ces trois utopies constituent-elles vraiment un triangle vertueux et un formidable accélérateur d'innovations comme certains discours médiatiques le laissent entendre ? Ou bien recèlent-elles des effets pervers, non désirés, qui vont à l'encontre du progrès social qu'elles assurent amener ? Examinons plus en détail leur mise en œuvre sur le terrain.

2. L'Ouverture (Open)

Dans le domaine de l'informatique, l'idéal de l'ouverture a été traduit par des principes ou des "grandes idées" qui ont guidé le design des systèmes d'information que Zittrain (2008 : 27) a bien analysés.

Le premier grand principe de design, le *verkeersbordvrij*, littéralement "*libre de presque tout panneau routier*", correspond à une expérimentation faite dans certains pays nordiques supprimant presque toute la signalisation routière. Le résultat fut contre-intuitif : les conducteurs faisaient plus attention aux piétons en utilisant éventuellement un signe de la main et se regardaient. Ce principe traduit l'idée que moins on impose des restrictions et de contrôle sur des inventions techniques, plus ces inventions résisteront face au rythme accéléré de changements et donc plus elles seront capables d'auto-évolution si nécessaire. Appliqué à l'invention de l'Internet et du *Wikipédia*, ce principe surpassa toutes les attentes. En effet, c'est bien cette ouverture à la fois initiale et génératrice qui permet ensuite de décliner possibilités, spécifications, fonctionnalités au moment de la concrétisation de l'artefact, car seul le processus de concrétisation réduirait progressivement, mais jamais totalement, la part d'abstraction des objets techniques en établissant "*une coordination opératoire entre les éléments de la machine et les éléments du milieu extérieur*" (Guchet, 2008 : 9). Autrement dit, les finitions et définitions des fonctionnalités et les modalités d'usages prescrites dans les objets techniques ne le sont ni totalement, ni antérieurement, par l'offre industrielle comme Lacroix, Tremblay et Mœglin l'ont démontré depuis longtemps (1992) et restent en partie aux mains des usagers (les contributeurs de *Wikipédia*) et des producteurs (les initiateurs porteurs du projet). La concrétisation, processus incessant et évolutif, se faisant dans une triple dialectique : lecteurs, contributeurs, organisateurs. Soulignons l'intérêt industriel, et donc économique, de ce design, que Simondon³ évoquait déjà cinquante ans plutôt. Pour lui en effet, le premier état d'un futur objet industriel restait celui de l'objet technique abstrait, principalement constitué de potentiels y compris pour les usages.

Le projet politique déjà décennal de l'ouverture des données publiques (Open Data) semble également adhérer à ce principe de spécification minimaliste. L'exigence politique porte sur la libéralisation des données publiques, sans restriction, et dans un format interopérable. A l'instar de l'information et la communication de Wiener qui devaient circuler sans entraves, l'Open Data ressuscite la métaphore des « autoroutes » de l'information, vulgarisée par le projet de *National Information Infrastructure* (NII) présenté par Al Gore en 1991 ; autoroutes sur lesquelles les données cette fois, sont appelées à circuler désormais librement et pour le progrès de la société.

La deuxième grande idée du design, issue de la première, est dans l'incitation à construire des technologies génératives. Selon Zittrain (2008 : 150), "*la générativité est la capacité d'un système*

² « Dans la Silicon Valley, le gouvernement est considéré comme lent, peuplé par des médiocres, infesté de règles obsolètes et d'inefficacité », Packer G., « Change the world » *The New Yorker*, 27 mai 2013. Visité le 9/05/2015, <http://www.newyorker.com/magazine/2013/05/27/change-the-world>.

³ Simondon, G. *Du mode d'existence des objets techniques*, Aubier, 1958.

à produire des changements non anticipés grâce à des contributions non filtrées/contrôlées d'une audience large et diverse".

Ces technologies génératives se caractérisent par le fait qu'elles sont "*ambitieuses dans leurs desseins mais modestes ou simples dans leur exécution*" (Zittrain 2008 : 134), ce qui sous-entend deux autres principes : (i) la procrastination, temporiser et ne résoudre les problèmes que lorsqu'ils surviennent et ; (ii) la robustesse, principe posé par Jon Postel, un des fondateurs de l'Internet Protocol (IP) qui l'a résumé ainsi : "*Soyez conservateur dans ce que vous faites : soyez libéral dans ce que vous acceptez des autres*"⁴.

Cet idéal de l'ouverture n'est pas né avec le développement de l'informatique de contenu (les logiciels). Il constitue l'un des piliers de la science que le sociologue fonctionnaliste américain Robert Merton (1973) formalisa en quatre impératifs constituant l'éthos de la science moderne : Communisme, Universalisme, Désintéressement et scepticisme organisé (**Organized Skepticism**), connu sous l'acronyme 'CUDOS'. L'idéal scientifique voudrait que toutes les données ayant servi à réaliser une étude soient accessibles dans un format ouvert pour permettre à d'autres scientifiques de s'en saisir. Le secret et la fermeture sont incompatibles avec l'éthos scientifique alors que l'ouverture et la publication des résultats sont ses impératifs. La communauté de la Digital Data Curation⁵ voudrait promouvoir au rang d'une exigence la production d'un '*data paper*' à côté de chaque publication scientifique (Tedds et Callaghan, 2014), sorte de compagnon de l'article scientifique décrivant les données utilisées et les conditions de collecte. Le ultime but rêvé étant l'interconnexion de toutes les données scientifiques, afin que les chercheurs puissent passer de la lecture d'un article scientifique, à la méthode de collecte des données utilisées et *vice versa*.

Au niveau sociétal, l'utopie de l'ouverture traduit l'hypothèse que l'exhibition permanente et planétaire de nos actes et la participation massive dans l'accomplissement d'une activité ont la capacité d'accélérer le rythme des innovations parce que, quelque part dans le monde, quelqu'un peut rectifier des informations ou données erronées en temps réel. Le succès de Wikipédia semble d'ailleurs accréditer cette utopie. Le revers de la médaille est que l'ouverture s'est accompagnée par le rétrécissement de la sphère privée et par la montée de l'exhibitionnisme, conséquences qu'Éric Schmidt, ancien PDG de la société Google admet sans détours : "*If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place*"⁶.

3. Le Gigantisme (Big)

L'immensité des données numériques disponibles a évidemment un impact considérable sur la conduite des affaires non seulement en science mais aussi dans la société. A l'horizon 2020, il est prévu⁷ que le volume de données numériques disponibles sur les réseaux atteigne 40 zetta-octets (un zetta-octets = 10²¹).

Les technologies de l'information et de la communication numériques (TICN) s'accompagnent de discours promotionnels qui visent à en faciliter la diffusion et l'appropriation. Dans ce discours, la métaphore joue un rôle important. Pushmann et Burgess (2014 : 1698) montrent que ces discours sont révélateurs des tentatives de naturalisation du phénomène des données, devenu quelque chose d'inéluctable, de "donné" justement ou de préexistant, tout en obscurcissant par là son aspect social et construit. Ainsi, les données sont-elles parfois présentées comme un phénomène naturel potentiellement dévastateur si sa force n'était pas contenue, d'où la surenchère dans l'emploi de substantifs tels "*océan, torrent, vague, déluge, tsunami*" accolés aux données. Mais à l'inverse, les données sont également présentées comme des ressources

⁴ Postel, J. "*Be conservative in what you do: be liberal in what you accept from others*", Transmission Control Protocol (TCP), janvier 1980 http://en.wikipedia.org/wiki/Robustness_principle

⁵ <http://www.dcc.ac.uk/digital-curation/why-preserve-digital-data>.

⁶ « Si vous avez quelque chose que vous ne voulez pas que personne sache, peut-être ne devriez-vous pas le faire du tout. », Google CEO on privacy VIDEO. http://www.huffingtonpost.com/2009/12/07/google-ceo-on-privacy-if_n_383105.html, consulté le 7 mai 2015

⁷ http://fr.wikipedia.org/wiki/Big_data consulté le 8 mai 2015

naturelles brutes et précieuses dont la valeur n'apparaîtra qu'au prix d'efforts importants de forage que seules les machines sont désormais en mesure de fournir. Comme les discours d'accompagnement liés à l'émergence de la fouille de données et de textes dans les années 1990, on retrouve les qualificatifs de "*pétrole, d'or, de minéraux ou de liquides précieux*".

Ce discours vulgarisateur et médiatique sur la toute-puissance des algorithmes du Big Data commencent à susciter une attitude de plus en plus critique, en particulier outre-Atlantique.

En remontant à l'origine du terme Big Data, Pushmann et Burgess (2014 : 1698) observent qu'il s'est réellement imposé dans le milieu du *business*. Au fond, le Big Data reste une affaire de commerce et de contrôle. Avare de concentrations et de monopoles, le courant ultralibéral a tout intérêt à maintenir l'illusion que le gigantisme constitue désormais la seule échelle valable pour innover au XXI^e siècle. Or, seules les très grandes firmes (GAFA) dotées de puissantes infrastructures informatiques robustes sont en mesure de capter et exploiter le Big Data. Grâce à leur position de *gatekeeper* de l'univers numérique, elles partagent le web en de multiples espaces numériques fermés et sous leur contrôle. Dans cet écosystème, les données des internautes, qu'elles ont captées gratuitement et cryptées dans des formats propriétaires deviennent des monnaies d'échange grâce auxquelles le commerce électronique s'impose comme mode principal de transactions commerciales.

Ce courant ultralibéral exploite d'ailleurs frénétiquement les données personnelles massives afin de générer des recettes publicitaires, au risque d'induire des usages immoraux ou illégaux. C'est ainsi que l'application *Google Education* "lisait" les courriels des jeunes usagers aux États-Unis afin de leur envoyer une publicité individualisée⁸.

Plus inquiétant encore, des start-up de l'économie numérique ont mis au point des applications aux noms évocateurs de *Whisper*⁹, *Creepy*¹⁰, *PleaseRobMe.com*¹¹, *The Girls Around Me*¹². Celles-ci agrègent automatiquement des données individuelles à partir de différentes plateformes de réseaux sociaux et autres dispositifs électroniques (cartes de transport urbain dotées de puces numériques, smartphones). Une fois agrégées, ces données géolocalisent des individus à un instant *t*, retracent l'historique de leurs déplacements et situent leurs domiciles (Ben Goldcare, 2014). Même lorsque ces applications affichent de soi-disant bonnes intentions, leur potentiel d'usages déviants reste non négligeable. *Samaritan*¹³, conçue pour prévenir les suicides en alertant la personne suivie sur *Twitter* de l'état moral de ses *followers* en fonction de l'emploi de certaines expressions, pouvait également traquer les individus vulnérables. Kosinski *et al.* (2013) ont montré qu'il était possible de prédire les principaux attributs des usagers de *Facebook* à partir des données de leurs comptes (orientations sexuelle, politique et religieuse, origines ethniques). Ils ont ainsi calculé l'orientation sexuelle d'un échantillon de 58 000 utilisateurs de *Facebook* alors que ceux-ci n'avaient pas fourni cette information, la jugeant trop personnelle. Il est en effet devenu beaucoup moins coûteux de surveiller les individus grâce aux dispositifs électroniques que d'engager des détectives humains (Bankston et Soltani, 2014).

Dans le secteur scientifique, le phénomène du Big Data peut engendrer de profonds changements. L'ère de la *e-Science* ou du Big Science implique une culture de collaboration à l'échelle internationale. S'il n'est pas question de nier les avancées permises par l'automatisation et la disponibilité des données massives au bénéfice de la science et de la société, il paraît tout aussi important de prendre conscience que la recherche à l'échelle du Big Data représente un changement épistémologique car non seulement elle déplace le focus sur ce qui est comptable ou mesurable, mais elle place l'étude des phénomènes sur une échelle macroscopique. Plusieurs auteurs avertissent des dangers de cette "data science". Jeffery Bowker (2013 : 170) craint ainsi

⁸ Jérôme Hourdeaux, Google accusé de ficher les élèves et étudiants, *Médiapart*, 25 mars 2014.

⁹ <http://www.theguardian.com/world/2014/oct/16/-sp-revealed-whisper-app-tracking-users>

¹⁰ <http://www.itproportal.com/2011/03/30/creepy-app-warns-end-privacy/>

¹¹ <http://mashable.com/2010/02/17/pleaserobme/>

¹² <http://www.cultofmac.com/157641/this-creepy-app-isnt-just-stalking-women-without-their-knowledge-its-a-wake-up-call-about-facebook-privacy/>

¹³ <http://www.theguardian.com/society/2014/nov/04/samaritans-twitter-app-mental-health-depression>

que "nos interactions avec le monde » ne se réduisent à "des épiphénomènes aux côtés des algorithmes de traitement de données. N'existerait alors que ce qui est quantifiable. Les collectivités et communautés qui ne sont pas quantifiées ou modélisées ne seraient préservées que de manière accidentelle. En tant que peuple, nous deviendrons, selon la formule mémorable d'Olga Kuchinskaya, nos propres données. Si vous n'êtes pas une donnée, vous n'existez pas ; et comme le malheureux Doc Daneeka dans *Catch 22*, peu importe le nombre de fois où vous clamez que vous êtes vivant."

L'abandon du savoir causal au profit d'une science des objets constitue un autre risque pour la science à l'échelle du Big Data car les scientifiques pourraient perdre à la fois la capacité d'expliquer les corrélations trouvées au sein des données massives par les algorithmes et celle de comprendre le pourquoi des phénomènes observés (Ibekwe-SanJuan, 2014).

L'engouement actuel pour une science orientée « données » semble opportunément balayer trois évidences que l'on croyait pourtant acquises au fil des siècles de recherche empirique :

1) aucune collection de données, aussi grande soit-elle ne peut être exhaustive. Il existerait par conséquent, toujours des « choses » en dehors de ces vastes collections des données. D'ailleurs, l'acte même de collecter des données n'est pas neutre car *"chaque acte visant à intégrer les données dans une archive est simultanément un acte d'occultation d'autres manières d'être, d'autres réalités. L'archive ne peut en principe contenir le monde en petit ; sa finitude signifie que la plupart des morceaux de réalité ne seront pas représentés. La question pour la théorie est quelles sont les formes d'exclusion et comment peut-on tirer des généralisations à leur propos ?"* (Bowker, 2014 : 1797)

2) le mode algorithmique de découverte imposé par l'investigation à l'échelle du Big Data ne conduit pas à des vérités irréfutables. Certaines études médiatisées, fondées sur le Big Data ont même produit des erreurs flagrantes. Les estimations erronées du pic de la grippe hivernale par *Google Flu Trends* (Auerbach, 2014) est symptomatique des errements de la « data science ». Crawford, Milner et Gray (2014) observaient à juste titre que *"les promesses célébrant le Big Data comme 'assez bonnes' pour prédire le comportement social ignorent une propriété fondamentale de la théorie sociale : les actions agrégées des individus ne peuvent pas, en elles-mêmes, illustrer la dynamique complexe qui engendre l'interaction sociale – l'ensemble de la société est plus grand que la somme de ses parties."* ;

3) les connaissances ne sont pas immanentes aux données, prêtes à être prélevées par la quantification et le calcul algorithmique. La connaissance nécessite une participation active des humains.

Thatcher (2014 : 1777) préconise que le discours sur le Big Data soit dépouillé *"de ses prétentions hyperboliques à atteindre la vérité absolue (...). Le Big Data doit être recontextualisé et ré-inséré dans le cadre des corps de savoirs socio-économique, culturel et politique."*

4. La Participation

La 'participation' mieux connue sous le terme anglais de *crowdsourcing*, renvoie la plupart du temps à la contribution d'un public nombreux versant des contributions sous forme de métadonnées ou de données élémentaires, sur une plateforme numérique dans le cadre de programmes de science participative ou d'initiatives citoyennes visant la création de biens communs numériques. Ces contributions du public sont encadrées. Les participants se trouvent alors dans une situation de subordination à une autorité détentrice de la plateforme qui valide les contributions. C'est la forme la plus répandue de participation en ligne et c'est dans ce sens que nous l'employons ici. On peut également parler de collaboration.

Troisième "grande idée" du design des technologies de l'information, la participation apparaît comme la troisième utopie de l'innovation au XXI^e siècle tant elle rencontre de succès et séduit de plus en plus d'acteurs inscrits dans des idéologies et logiques différentes.

Elle traduit l'idée d'une confiance accordée aux amateurs et au grand public dans la production de biens et de services numériques. Souvent cité comme témoin de la réussite du modèle participatif, le succès de l'expérience du *Wikipédia* s'expliquerait également par l'inscription du principe de *verkeersbordvrij* (spécification minimaliste) dans son design. L'objectif de ses inventeurs était de construire une encyclopédie dotée de peu règles sinon celles de maintenir un point de vue neutre ; d'éliminer des contenus qui enfreignent les lois du copyright ; d'ignorer celles qui interfèrent avec l'objectif premier de construire l'encyclopédie et enfin de supprimer les *gatekeepers* (Zittrain, 2008 : 109).

Dans son célèbre essai '*The Cathedral and The Bazaar*', référence de la philosophie 'open source', Raymond (1999) oppose deux approches de *software design*. La cathédrale, classique en économie politique, se caractérise par un excès de centralisation et d'hierarchie, elle définit soigneusement le rôle et la place de chacun et les fait s'enchaîner. La deuxième, plus coopérative, a une structure de type réseau au sein duquel les rôles ne sont ni clairement définis, ni même assignés. Elle réunit des gens ayant éventuellement des approches et des agendas différents au sein d'un même projet où règne une sorte de désordre organisé. Ce deuxième modèle du Grand Bazar s'est révélé propice à l'innovation sociale et bénéfique à l'innovation industrielle car il provoque d'une part un brassage qui fait barrière à la routine et de l'autre, des associations improbables ou opportunités hasardeuses qui ne sauraient résulter de calculs ou de modes opératoires conceptualisés.

Les soirées hackathon autour des projets de l'Open Data illustrent ce modèle de développement. Notons de plus que le double accès à ce processus d'innovation d'abord par des professionnels puis par des amateurs au XX^e siècle, tel que Leadbeater et Miller l'analysent (Demos, 2004) confirme à la fois le succès et la faisabilité de la mise en œuvre de ce principe.

Cependant, sur le terrain, le modèle de développement des logiciels libres n'est pas aussi horizontal ni aussi hasardeux que Raymond le laissait entendre. Le développement des logiciels libres et pas seulement, se caractérise quasiment toujours par un « noyau dur » de développeurs déterminant les grandes orientations du projet et autour duquel gravitent une communauté d'utilisateurs et de contributeurs occasionnels comme l'ont montré Leadbeater et Miller en parlant de nébuleuse d'innovation. Il s'agit donc d'une autre utopie qui doit composer avec des paramètres de type organisationnels, nécessaires à l'innovation.

Le modèle de la participation et le *crowdsourcing* déjà actés comme modèles économiques (Bouquillion, Matthews, 2010) sont en passe de devenir également des modes de gestion usuels des affaires publiques, à côté des algorithmes informatiques, sur un principe issu des industries créatives. Le *crowdsourcing* constitue une alternative sérieuse et complémentaire à l'automatisation et apparaît de plus en plus déterminant dans certains domaines d'activités, confrontés à des problèmes que les algorithmes du Big Data sont (encore) incapables de résoudre. Pour le courant libertaire, cette participation du public dans la co-construction des connaissances est un gage que les biens et les services ainsi produits resteront ouverts et dans le domaine public.

La recherche en génomique fait appel de plus en plus au *crowdsourcing* pour annoter et reconnaître des motifs. Les algorithmes travaillent ensuite sur ces annotations faites par les volontaires afin de chercher des corrélations en toutes directions. L'enjeu n'est rien moins que le brevetage du vivant¹⁴. Le premier à trouver un séquençage ou un lien entre un gène et une maladie peut le breveter, d'où une course entre recherche publique et privée, les acteurs publics tentant de mobiliser très largement les citoyens bénévoles non pas pour révolutionner la génomique, mais pour que les applications potentielles de ces recherches restent dans le domaine public.

Selon Atal Butte (2014), l'industrie pharmaceutique est en crise à cause du coût élevé de la fabrication de nouveaux médicaments. Dans le même temps, le volume de données biomédicales en accès libre dépasse les capacités de traitement des chercheurs et scientifiques du domaine. Désormais, des entreprises se mettent à commercialiser des tests de laboratoires en ligne¹⁵, à

¹⁴ https://fr.wikipedia.org/wiki/Brevetage_du_vivant

¹⁵ <https://www.assaydepot.com/about>

partir d'échantillons jetés par les hôpitaux américains par manque de capacités de stockage. Pour continuer à disposer de nouveaux médicaments à un prix abordable, il pourrait bien échoir au grand public d'explorer ces masses de données pour tenter de découvrir de nouvelles molécules ou des nouveaux diagnostics. Ainsi, la participation du public est déterminante dans ce secteur.

En astronomie, le projet SDSS a donné lieu au projet de science participative *Galaxy Zoo*¹⁶ qui met à contribution les amateurs afin de classifier les images d'objets célestes collectées par le télescope SDSS. Lancé en 2007, il a suscité l'aide de plus de 150 000 participants qui ont conduit les scientifiques à détecter plus de 230 millions d'objets célestes.

De même, le projet *eBird*¹⁷ mené par l'Université de Cornell sollicite la participation des amateurs dans l'immense tâche de recensement et de description de toutes les espèces d'oiseaux, mission impossible pour les scientifiques au regard de leur temps et moyens. Cette initiative a permis de recueillir jusqu'à 160 millions observations venant de plus de 1000 observateurs et couvrant plus de 95% des espèces d'oiseaux. Nous sommes bien loin du temps où l'expert patenté ou le scientifique étaient les seuls habilités à accomplir certaines tâches et surtout à s'y intéresser.

Des centaines d'autres projets de science participative sont recensés (Heaton & Proulx, 2015).

Plusieurs observateurs sont persuadés que les connaissances, les biens et les services seront de plus en plus co-construits par des amateurs qui sont de plus en plus jeunes (Butte, 2014).

Open, Big, Participation : un triangle vertueux ou vicieux ?

Si les deux idéologies – libertaire et ultralibérale, sous-jacentes à ces trois utopies de l'innovation paraissent de prime abord contradictoires, il n'en demeure pas moins que sur le terrain de l'innovation, leurs modalités de mise en œuvre se conjuguent. L'ouverture implique la participation tout en accroissant le phénomène du gigantisme, tandis que la participation de la foule contribue à alimenter les algorithmes du Big Data en corpus d'apprentissage qui leur permet de s'affiner et d'améliorer leurs performances.

L'image d'un triangle semble mieux rendre compte des interactions entre ces trois utopies.

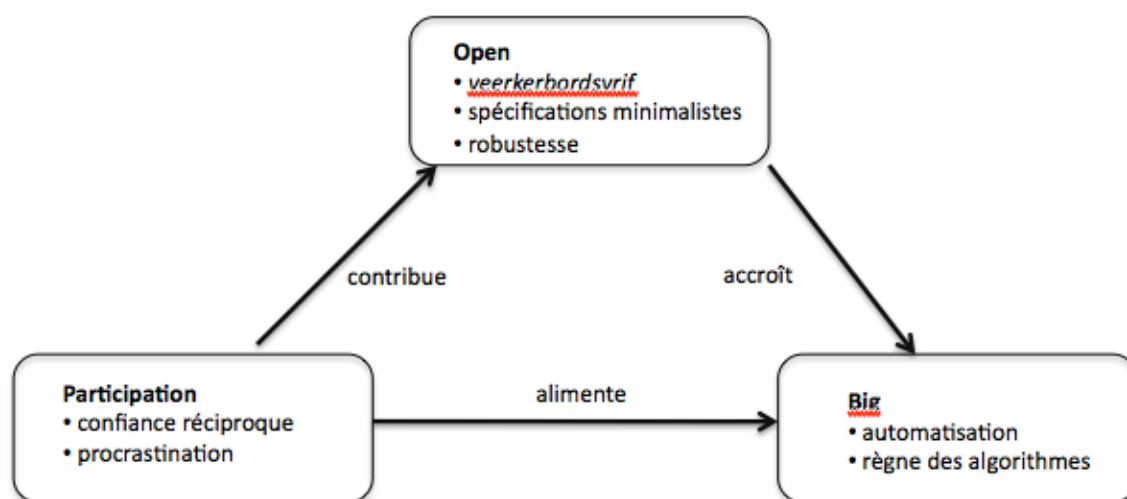


Figure 1. Interactions entre les trois utopies : ouverture, participation et gigantisme.

Cette alliance qui semble contre nature entre libertaires et ultralibéraux n'est pourtant pas nouvelle. Dans *"Aux Sources de l'utopie numérique. De la contre-culture à la cyberculture, Stewart Brand, un homme d'influence"*, Fred Turner (2006) montre qu'elle était déjà à l'œuvre dans la conception même de l'Internet, fruit de la rencontre de la contre-culture et des néo-

¹⁶ <http://www.galaxyzoo.org/>

¹⁷ <http://ebird.org/content/ebird/>

communalistes américains dans les années 1960, porteurs d'une idéologie libertaire avec l'idéologie libérale et capitaliste des pionniers de la Silicon Valley. Ces deux mouvements s'unissaient dans leur affirmation de la liberté individuelle créative, le rejet de la politique considérée comme étant le problème, leur amour pour la technologie numérique vue comme moyen d'émanciper l'homme et de changer la société.

Dans "L'Age de faire", Lallement (2015) constate également qu'"un même appétit de liberté fonde la doctrine anarchiste et celle d'un capitalisme plus débridé". L'idéologie libérale supplante l'idéologie libertaire lorsque l'attrait du marché devient trop fort. Cependant, ces deux courants idéologiques – libertaire et ultralibéral, captent les fruits du *crowdsourcing*, selon des objectifs opposés.

Ces trois utopies constituent-elles pour autant un triangle vertueux pour la science et la société ? Conjuguées, elles participent à l'émergence de ces technologies disruptives qui érodent peu à peu les monopoles des expertises dans plusieurs secteurs. L'entrée de la plateforme *AirBnB* dans le secteur de l'hôtellerie et de la société *Über* dans le secteur des transports urbains illustre la capacité qu'ont ces trois utopies à déstructurer l'équilibre entre acteurs économiques traditionnels d'un secteur par l'accroissement de l'entropie (désordre et incertitude). Pour le moment, seul le courant régalien est en mesure de juguler les effets de cette entropie par la recherche d'équilibre ou d'homéostasie ; les récentes législations pour le droit à l'oubli numérique¹⁸ par exemple ou d'autres législations visant à contraindre les sociétés du web à expliciter les termes et usages de leur collecte passive des données individuelles.

Fondée sur l'écosystème des données, l'économie numérique a besoin d'un public massif qui est devenu tour à tour la matière première qui nourrit gratuitement cette économie en données personnelles sans contrepartie ; une main d'œuvre gratuite *via* le *crowdsourcing* et les contributions en ligne ; la cible d'un marketing profilé grâce aux traces numériques fournies par ce même public et analysées par les algorithmes du Big Data.

Cependant, les participants ne sont pas devenus pour autant les nouveaux esclaves de l'économie numérique et du courant ultralibéral. Ceux qui participent à une initiative de création de biens numériques le font soit par idéologie libertaire (pour protéger les « communs de la connaissance » contre la logique ultralibérale marchande), soit par pulsion épistémique (par désir de partager leurs connaissances et obtenir une reconnaissance auprès de leurs pairs).

Cette capacité qu'a le numérique à être à la fois source d'émancipation sociale et politique pour les citoyens, mais en même temps, source d'aliénation ou de soumission, est qualifiée de "*paradoxical empowerment*" par Proulx et Heaton (2015). Ce paradoxe engendre à la fois des "technologies génératives" et des évolutions sociétales inattendues qui seront à étudier attentivement.

Bibliographie

ANDREJEVIC (M.), *The Big Data Divide*, dans *International Journal of Communication*, vol. 8 2014, pp. 1673–1689.

AUERBACH (D.), The Mystery of the Exploding Tongue. How reliable is Google Flu Trends?, *Slate.com*, 19 mars 2014. Consulté le 8 mars 2015.

http://www.slate.com/articles/technology/bitwise/2014/03/google_flu_trends_reliability_a_new_study_questions_its_methods.html.

BANKSON (K. S.) et SOLTANI (A.), Tiny constables and the cost of surveillance: Making cents out of United States v. Jones. *Yale Law Journal*, 123, 9 janvier 2013. <http://yalelawjournal.org/forum/tiny-constables-and-the-cost-of-surveillance-making-cents-out-of-united-states-v-jones>

BOUQUILLON (Ph.) et MATTHEWS (J.), *Le Web collaboratif : mutations des industries de la culture et de la communication*, Grenoble : P.U.G., 2010.

BOWKER (J.), *The Theory/Data Thing. Commentary*, dans *International Journal of*

¹⁸ Martin Untersinger, Internet : du devoir de mémoire au « droit à l'oubli », *Le Monde*, 24.09.2014

Communication, vol 8, 2014, pp. 1795–1799.

BOWKER (G. C.), *Data flakes: An afterword to “raw data” is an oxymoron* dans L. Gitelman (sous dir.), *“Raw data” is an oxymoron*, Cambridge, MA: MIT Press, 2013, pp. 167–171.

BRETON (Ph.) et PROULX (S.), *L'Explosion de la communication : introduction aux théories et aux pratiques de la communication*, Paris, La Découverte, 2005.

BUSCH (L.), *A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large-Scale Data Sets*, dans *International Journal of Communication* vol 8, 2014, pp. 1727–1744.

CRAWFORD (K.), MILNER (K.), GRAY (M.L.), Critiquing Big Data: Politics, Ethics, Epistemology, *International Journal of Communication*, 8, 2014, pp. 1663–1672. <http://ijoc.org>.

CRAWFORD (K.) (2013 1er avril), “The hidden biases in Big Data”. Consulté le 5 janvier 2015 http://blogs.hbr.org/cs/2013/04/the_hidden_biases_in_big_data.html.

GOLDACRE (B.) (2014 5 déc.), “When data gets creepy: the secrets we don’t realise we’re giving away”, *The Guardian Technology*. <http://www.theguardian.com/technology/2014/dec/05/when-data-gets-creepy-secrets-were-giving-away>.

GUCHET (X.), Évolution technique et objectivité technique chez Leroi-Gourhan et Simondon, *Appareil*, 2008. Consulté le 05 janvier 2015. <http://appareil.revues.org/580>.

IBEKWE-SANJUAN (F.), Big Data, Big machines, Big Science : vers une société sans sujet et sans causalité ?, 19^{ème} Congrès de la SFSIC, Toulon 4-6 juin 2014, 9 pages. <http://sfsic2014.sciencesconf.org/browse/session?sessionid=5409>.

KOSINSKI (M.), STILLWELL (D.), GRAEPEL (T.), *Private traits and attributes are predictable from digital records of human behavior*, dans *Proceedings of the National Academy of Sciences (PNAS)*, 9/04/2013, vol. 110, n°. 15, pp. 5802–5805. www.pnas.org/cgi/doi/10.1073/pnas.1218772110.

HEATON (L.), et PROULX (S.), *Paradoxical Empowerment: Immaterial Labor Translated in a Web of Affective Connections* dans *The Information Society: An International Journal*, 2015, 31:1, pp. 28-35, <http://dx.doi.org/10.1080/01972243.2015.977629>.

JANIER (J.), *Who Owns the Future?*, Simon & Schuster, 2014.

LACROIX (J-G.), TREMBLAY (G.), MOEGLIN (P.), (1992), *Usages de la notion d’usages* dans *Les nouveaux espaces de l’information et de la communication*, Inforcom, 8^{ème} Congrès de la SFSIC, Lille, pp. 241-248.

MERTON (R. K.), The Normative Structure of Science, dans R.K. Merton, *The Sociology of science. Theoretical and empirical investigations*, University of Chicago Press, 1973, pp. 267-278.

MOROZOV (E.), *To Save Everything, Click Here: The Folly of Technological Solutionism*, PublicAffairs, 2014.

PROULX (S.), *Usages participatifs des technologies et désir d’émancipation : une articulation fragile et paradoxale*. A paraître dans ‘Communiquer. Revue internationale de communication sociale et publique’, numéro sur les « Perspectives en communication », Montréal, hiver 2015, pp. 35-45.

PUSHMANN (C.) et BURGESS (J.), Metaphors of Big Data, *International Journal of Communication*, vol. 8, 2014, pp.1690–1709

SHIRKY (C.), *Power Laws, Weblogs, and Inequality*, 2003 http://www.shirky.com/writings/powerlaw_weblog.html.

THATCHER (J.), Living on Fumes: Digital Footprints, Data Fumes, and the Limitations of Spatial Big Data, *International Journal of Communication*, vol. 8, 2014, pp.1765–1783

TEDDS (J.) et CALLAGHAN (S.), *Guidelines on recommending data repositories as partners in data publication* dans 9th International Digital Data Curation Conference, 24-27 février 2014, San Francisco, <http://www.dcc.ac.uk/events/idcc14/day-two-papers>.

TURNER (F.), *Aux sources de l’utopie numérique : De la contre-culture à la cyberculture*, Stewart Brand, un homme d’influence. C&F Editions, 2012.

ZITTRAIN (J.), *The Future of the Internet*, Penguin Books, 2008. <http://futureoftheinternet.org/files/2013/06/ZittrainTheFutureoftheInternet.pdf>.