

FRANCESCA MUSIANI

Nains sans géants

Architecture décentralisée
et services Internet

Préface de Geoffrey C. Bowker

Prix Informatique et Libertés 2013



Sciences sociales

Presses des Mines

Francesca Musiani, *Nains sans géants. Architecture décentralisée et services Internet*, Paris, Presses des Mines, Collection Sciences sociales, 2^e édition, 2015.

© Presses des MINES - TRANSVALOR, 2015

60, boulevard Saint-Michel - 75272 Paris Cedex 06 - France

presses@mines-paristech.fr

www.pressesdesmines.com

ISBN : 978-2-35671-146-5

© Photo de couverture : Danièle Akrich

Dépôt légal : 2015

Achévé d'imprimer en 2015 (Paris)

Tous droits de reproduction, de traduction, d'adaptation et d'exécution réservés pour tous les pays.

Nains sans géants

Collection Sciences sociales

Responsable de la collection : Cécile Méadel
Centre de sociologie de l'innovation (www.csi.ensmp.fr)

- L. Doganova, *Valoriser la science, Les partenariats des start-up technologiques.*
- F. Granjon, *Reconnaissance et usages d'Internet.*
- F. Massit-Folléa, C. Méadel, L. Monnoyer-Smith, *Normative Experience in Internet Politics.*
- M. Calvez avec la collaboration de S. Leduc, *Des environnements à risques.*
- G. Teil, *Le Vin et l'Environnement.*
- J. Denouël, F. Granjon (dir.), *Communiquer à l'ère numérique, Regards croisés sur la sociologie des usages.*
- A. Mallard, *Petit dans le marché, Une sociologie de la Très Petite Entreprise.*
- M. Akrich, T. Barthe, F. Muniesa, P. Mustar (eds.), *Débordements, Mélanges offerts à Michel Callon.*
- M. Akrich, Y. Barthe, C. Rémy, *Sur la piste environnementale.*
- C. Lemieux, *Un président élu par les médias ?*
- C. Lemieux, *La Sociologie sur le vif.*
- M. Armatte, *La Science économique comme ingénierie.*
- J. Denis et D. Pontille, *Petite sociologie de la signalétique, Les coulisses des panneaux du métro.*
- A. Mol, *Ce que soigner veut dire, Repenser le libre choix du patient.*
- M. Akrich, C. Méadel, V. Rabeharisoa, *Se mobiliser pour la santé. Les associations de patients témoignent.*
- M. Akrich, J. Nunes, F. Paterson, V. Rabeharisoa (eds), *The Dynamics of Patient Organizations in Europe.*
- M. Mort, C. Milligan, C. Roberts, I. Moser (ed.), *Ageing, Technology and Home Care.*
- M. Akrich, M. Callon et B. Latour, *Sociologie de la traduction. Textes fondateurs.*
- A. Desrosières, *Pour une sociologie historique de la quantification, L'Argument statistique I.*
- A. Desrosières, *Gouverner par les nombres, L'Argument statistique II.*
- A. Savoye, F. Cardoni (coord.), *Frédéric Le Play, parcours, audience, héritage.*
- F. Audren, A. Savoye (eds), *La Naissance de l'ingénieur social.*
- A. de Saint Laurent-Kogan, J.-L. Metzger (dir.), *Où va le travail à l'ère du numérique ?*
- B. Latour, *Chroniques d'un amateur de sciences.*
- V. Rabeharisoa, M. Callon, *Le Pouvoir des malades.*

Nains sans géants

Architecture décentralisée
et services Internet

Francesca Musiani

Préface de Geoffrey C. Bowker

2^e édition
Prix Informatique et Libertés 2013

Préface

Faire l'histoire de l'imprimerie conduit aisément à tenir un récit de l'inéluctable : dès lors que l'histoire s'est déroulée de cette façon, c'est donc ainsi qu'il fallait qu'elle se déroulât. Tous ses éléments semblent avoir éclos par une même nécessité : le concept d'auteur, l'ossature des notes, les bibliographies, le cadre juridique du droit d'auteur. On en arrive facilement à raconter que cette histoire avait à faire avec la démocratisation, traduisant la parole de Dieu dans les langues vernaculaires, rajoutant au tiers-État, un quatrième et ouvrant la voie à un âge d'or de l'humanisme.

Que rien de tout cela n'ait été obligé – et qu'une grande partie de ce récit soit inexact – n'atténue pas la force de cette mythologie. Des récits moins optimistes commencent à montrer que l'imprimerie peut aussi être appréhendée comme instrument essentiel d'une bureaucratie totalisante (suivant l'argument de Michael Clanchy), ou encore que le format imprimé prime désormais quantitativement sur le livre imprimé (comme le soutient Lisa Gitelman).

Nous vivons actuellement une conjoncture historique chargée d'implications sociales, politiques et culturelles tout aussi importantes : le développement de l'Internet. Et déjà l'histoire et les histoires, à l'intention des générations futures, sont racontées : de l'effet « inévitablement démocratique » des médias sociaux, à l'inéluctabilité de la concentration des médias. Deux récits, faux et rassurants.

Dans ce formidable ouvrage, Francesca Musiani pose la question – en temps réel – de ce que signifierait penser l'Internet différemment. S'appuyant sur une ethnographie riche et une solide approche théorique, elle raconte magnifiquement bien les tentatives visant à créer des technologies pair-à-pair de moteur de recherche, de stockage de données ou de streaming vidéo.

Si la gouvernance de l'Internet est l'une des questions sociotechniques fondamentales de notre époque, nous devons être en mesure d'explorer à la fois les choix que nous avons faits et les routes que nous n'avons pas empruntées. Pour affronter ce problème politique crucial de notre temps, il nous faut comprendre ce qui est en jeu d'un point de vue social et comment la technologie peut se transformer.

Lorsque de nouvelles infrastructures informationnelles sont forgées, leur stabilisation en des formes rigides peut prendre des centaines d'années (l'arc décrit par l'imprimerie, du siècle des Lumières au système universitaire, en est un exemple). Tout ce qui concerne notre relation sociale avec les données, l'information et la connaissance est en cause.

Le grand mérite de ce livre est de fournir un instrument solide qui permet de penser ces questions. Non seulement de penser, dans une perspective historique, la façon dont «il aurait pu en être autrement», mais aussi d'explorer comment «il pourrait en être autrement».

Geoffrey C. Bowker
Professeur à l'université de Californie, Irvine

Pour mes parents

Remerciements

Mon premier remerciement reste pour Cécile Méadel – qui a dirigé la thèse dont cet ouvrage est issu.

Le deuxième est pour la Commission Nationale de l'Informatique et des Libertés (CNIL) qui m'a fait l'honneur de m'attribuer son Prix de thèse en 2013, en rendant possible cette deuxième édition. Merci à mes deux rapporteurs pour le Prix, Daniel Le Métayer et Fabrice Rochelandet, pour leurs remarques et suggestions, et à Sophie Vulliet-Tavernier et Marie Leroux pour leur suivi et coordination.

Le Prix Informatique et Libertés a ouvert la voie à ma nomination, en mai 2014, au sein de la Commission «droits et libertés à l'âge du numérique» de l'Assemblée nationale. Un double merci, donc, à la CNIL pour contribuer à rendre cette recherche significative dans l'espace public aussi bien que dans l'espace académique.

Valérie Beaudouin, Geoffrey Bowker, Massimiano Bucchi, Alexandre Mallard et Laurence Monnoyer-Smith ont commenté mon travail alors qu'il revêtait encore ses habits de thèse. La préface que Geoff m'a fait l'honneur d'écrire il y a un an n'a jamais été plus pertinente qu'en ces moments de révélations Snowden et débats autour du droit à l'oubli. Merci également à mes collègues de l'université de Georgetown et du *Berkman Center for Internet and Society* de l'université de Harvard, ainsi qu'à mes «nouveaux» collègues de l'Institut des sciences de la communication du CNRS.

Le Centre de Sociologie de l'Innovation de MINES ParisTech a été, pendant six ans, un environnement de travail idéal, ainsi que deux projets, financés par l'Agence Nationale de la Recherche : Vox Internet (2008-2010) et ADAM (Architectures distribuées et applications multimédias, 2010-2014).

Cette recherche n'existerait pas sans la disponibilité et l'enthousiasme des chercheurs, développeurs et entrepreneurs qui travaillent chaque jour avec les «nains» de l'Internet. Je leur suis reconnaissante pour le temps qu'ils m'ont consacré, et j'espère qu'ils trouveront ce travail utile. Merci aux spécialistes du P2P qui ont répondu à mes questions dans la phase exploratoire de ce travail, en particulier à Fabrice Le Fessant.

Plusieurs collègues, amis et proches ont, par leur soutien et leurs relectures, grandement contribué à cet ouvrage ainsi qu'à sa deuxième édition. Vous savez qui vous êtes : merci, du fond du cœur, pour m'accompagner au quotidien dans ce parcours à la fois professionnel et très personnel qu'est la recherche.

Avant-propos à la deuxième édition

«Je crois qu'en ce moment historique, le plus grand danger pour nos libertés et notre mode de vie dérive des pouvoirs omniscients des États, et du fait qu'ils ne soient contraints que par des documents politiques,» déclarait Edward Snowden le 10 juin 2013¹, quelques jours après avoir été à l'origine des premières fuites de documents *top secret* révélant le programme de surveillance systématique des réseaux mis en œuvre par la *National Security Agency*. Je me souviens de comment, pendant ces jours de tourmente politique, j'étais partagée entre l'admiration pour le courage de cet informaticien – employé par un prestataire de services de l'agence de renseignements américaine – et l'impression d'avoir sous les yeux les preuves incontournables d'une hypothèse qui, bien que maintes fois formulée, semblait en même temps relever de la science-fiction. Mais surtout, je me rappelle d'une autre pensée qui s'est alors manifestée : «Voilà ce qui change la donne pour les questions que je traite dans mon travail – j'aurais voulu et j'aurais dû en traiter dans le livre.» La première édition du présent ouvrage avait, en effet, été publiée juste quelques jours avant, le 6 juin 2013. Je me réjouis que l'attribution du Prix Informatique et Libertés 2013 à mes recherches me donne aujourd'hui l'occasion de combler cette lacune, au moyen de cette deuxième édition.

Le Prix n'est, en effet, pas seulement un honneur. Il est aussi l'occasion pour moi de revenir sur presque six ans de travail autour des implications sociales, politiques, juridiques et économiques du «choix de la décentralisation» dans l'architecture des services Internet, en développant des aspects qui – bien que déjà présents dans l'ouvrage initialement publié – doivent aujourd'hui être placés tout particulièrement à l'avant-plan. La protection de la vie privée, les différentes manières de la définir, la variété d'instruments qui peuvent l'assurer – par la politique, par la technique, par le droit – face à la surveillance des réseaux : ces questions sont étroitement liées aux choix de conception et de développement de l'architecture technique qui sous-tend nos services Internet, et demandent, aujourd'hui plus que jamais, l'attention des pouvoirs publics, de la société civile organisée, et des citoyens.

La deuxième édition de cet ouvrage est aussi une occasion de contribuer à l'évolution de la conception que le grand public a de la gouvernance d'Internet, en tant que champ d'études et de pratiques. Il fut en effet un temps, au début des années 2000, où se définir comme spécialiste de questions de gouvernance

¹ http://www.washingtonpost.com/world/national-security/code-name-verax-snowden-in-exchanges-with-post-reporter-made-clear-he-knew-risks/2013/06/09/c9a25b54-d14c-11e2-9f1a-1a7cdee20287_print.html

de l'Internet suscitait l'indifférence sinon la perplexité auprès des citoyens. Cette étiquette, «gouvernance de l'Internet», semblait en effet renvoyer à un monde étrange et étranger à notre vie quotidienne: un monde d'institutions onusiennes et intergouvernementales ou encore de standardisation technique, peuplées d'acronymes indigestes, s'occupant de questions et de jeux de pouvoir éloignés des préoccupations des utilisateurs.

Les temps ont changé. Pas pour la pertinence des questions auxquelles renvoie l'étiquette «gouvernance de l'Internet»: celles-ci n'ont jamais cessé d'être fondamentales. Mais parce que, du fait du nombre de problèmes soulevés quotidiennement – de l'omniprésente surveillance des réseaux au débat sur l'obligation des moteurs de recherche d'assurer le «droit à l'oubli» à toute personne qui en ferait la demande – on commence désormais à comprendre que la gouvernance d'Internet se construit également, et peut-être principalement, au fil des dispositifs et des pratiques qui façonnent le «réseau des réseaux» jour après jour.

L'Internet, autrefois conçu et utilisé dans une poignée de laboratoires informatiques publics et privés américains, intéresse désormais l'ensemble des activités humaines et est rendu accessible sur une variété de supports croissants, s'appêtant même à «coloniser» en masse les objets de notre vie quotidienne. Cette omniprésence discrète suscite des questions relatives à la vie privée, à la formation des collectifs, à la reconfiguration des marchés; elle se traduit notamment par la création de nouveaux instruments techno-juridiques pouvant assurer la protection des libertés individuelles et garantir l'existence ainsi que le progrès des arènes démocratiques.

Géants et concentration : une tendance anticipée

Dans les années 1970, l'Internet est un collectif d'ordinateurs, décentralisé et peu étendu. L'introduction à grande échelle des ordinateurs personnels et domestiques qui s'ensuit se développe en parallèle à un «optimisme technologique» représenté par le *manifesto* de John Perry Barlow en 1996, la célèbre Déclaration d'indépendance du cyberspace. Barlow décrit à cette occasion une utopie numérique et anarchiste, au sein de laquelle les citoyens du Net ou *Netizens* s'organiseraient autour de formes d'auto-régulation autonomes et ne devraient aux institutions d'antan que la curiosité académique qu'on doit aux fossiles. «Au nom du futur, je vous demande de nous laisser tranquilles. Vous n'êtes pas les bienvenus et vous n'avez aucune souveraineté où l'on se réunit,» déclara Barlow.

Presque vingt ans de développements, évolutions et contre-évolutions d'une portée considérable ont désormais rendu l'Internet quelque chose de très différent de cette vision (il est d'ailleurs discutable que celle-ci ait jamais été un portrait fidèle de l'Internet en pratique). L'Internet qu'on connaît aujourd'hui est un «réseau de réseaux» où un pourcentage impressionnant de l'organisation et de la circulation des flux et de l'information passe par un nombre limité de grandes entreprises. Informations et flux qui sont donc profondément susceptibles d'être influencés par ces entreprises et par les institutions avec qui elles entretiennent, maintes fois, des «liaisons dangereuses». Si Google, par exemple, constitue désormais un quart du trafic Internet nord-américain, l'influence globale du géant de l'information en ligne n'a été que trop visible lorsque la panne d'août 2013 a causé une diminution de quarante pour cent du trafic mondial².

Il s'agit là d'une «convergence» ou tendance à la concentration qui avait été anticipée par nombre d'ingénieurs et développeurs pionniers de l'Internet. Paul Baran, l'un des deux inventeurs des réseaux à commutation de paquets qui auraient donné naissance à l'Internet, prédisait depuis 1967 l'avènement d'un service («*utility*») informatique centralisé, qui aurait offert de la capacité de calcul de la même manière que les fournisseurs d'énergie nous font parvenir l'électricité. Modèle dont on peut aisément retrouver l'équivalent dans les empires informationnels d'Amazon, Google, Microsoft, et d'autres entreprises qui s'inspirent du paradigme du *cloud computing*. Et comme Paul Baran avait prédit, ces services nous offrent, fondamentalement, une impression de confort et d'immédiateté au détriment de la protection de la vie privée.

Les utilisateurs de l'Internet et des services qui le peuplent se soumettent régulièrement, en effet, à des politiques de confidentialité et conditions d'utilisation qui permettent aux entreprises de partager leurs données personnelles avec des institutions gouvernementales ou intergouvernementales, ou avec des entreprises tierces, à des fins de marketing. Parfois, des lois qui précèdent l'avènement du Web déterminent le cadre juridique qui lui est applicable: c'est le cas des États-Unis, où l'*Electronic Communications Privacy Act* – dont on a souvent invoqué une modernisation – permet aux agences étatiques d'obtenir sans un mandat les données personnelles que le citoyen confie à des tiers, y compris les données de présence collectées et agrégées de manière passive depuis les téléphones portables et les courriels.

² Sky News, «Google Outage: Internet Traffic Plunges 40%», <http://news.sky.com/story/1129847/google-outage-internet-traffic-plunges-40-percent>

Les révélations Snowden et la « contre-surveillance par la décentralisation »

En juin 2013, une bombe explose, sous la forme des révélations de l'informaticien Edward Snowden, prestataire de service extérieur pour la *National Security Agency* (NSA), le principal service de renseignement américain. Snowden montre au monde – ou plutôt, fait éclater au grand jour, documents *top secret* à la main – que ces imposants silos d'informations permettent aux agences de renseignement de cibler avec précision leurs opérations en étant capables d'être à l'écoute d'une grande partie de la population mondiale.

L'une de ces révélations, en particulier, suscite l'indignation : la NSA serait en train d'intercepter discrètement les connexions entre les puissants centres de données appartenant à Google et Yahoo!, permettant à l'agence de renseignements de collecter les données personnelles des utilisateurs alors qu'elles circulent sur les réseaux de ces entreprises. Les ingénieurs de Google réagissent en promettant le chiffrement de ces connexions afin d'éviter « par la technique » toute intrusion future ; Yahoo! fait une déclaration similaire. Les hauts dirigeants de Microsoft, quant à eux, font savoir non seulement qu'ils mettront en œuvre des mesures de sécurité semblables, mais qu'ils sont sur le point d'ouvrir des « centres de transparence » qui passeront au crible la plus grande partie du code source de leurs logiciels, à la recherche de *back doors* installés à leur insu par des agents liés au gouvernement. Sur le site Reformgovernmentsurveillance.com, huit acteurs majeurs de l'industrie IT américaine – d'habitude en compétition farouche les uns avec les autres – s'unissent pour demander un renforcement de la transparence des activités gouvernementales, et une modernisation des lois relatives à la surveillance.

Les déclarations enflammées de Mark Zuckerberg ou de Marissa Mayer, plaidant leur engagement en faveur de la vie privée et une plus grande transparence, ne suffisent cependant pas à restaurer la confiance des utilisateurs/citoyens envers les géants américains du *cloud*. Cette confiance est ultérieurement mise à l'épreuve lorsque des nouvelles révélations s'accumulent, explorant les « liaisons dangereuses » (et formelles) entre nombre d'entreprises IT et la NSA ; la surveillance massive de contenus qui se passe à leur entrée et leur sortie des États-Unis, avec d'importantes implications pour la surveillance internationale ; les tentatives délibérées de la part du service de renseignement américain d'affaiblir dès leur naissance les standards cryptographiques. En novembre 2013, le *New York Times* publie un document qui liste spécifiquement, parmi les missions de la NSA, celle d'« influencer le marché mondial des technologies d'encryptage » au moyen de relations commerciales avec des entreprises qui développent et implémentent des produits de sécurité informatique.

Dans ce contexte, la solution explorée par plusieurs développeurs devient celle de travailler à un Internet qui soit plus semblable à celui des origines : moins centralisé, plus distribué. Un projet qui ouvre des opportunités et pose en même temps des défis de taille : si, d'un côté, les solutions techniques décentralisées promettent à l'utilisateur un contrôle plus fin sur ses données, une meilleure protection de ces données avant qu'elles ne « sortent » de son ordinateur, elles posent en parallèle des problèmes de prise en main des dispositifs, d'utilisation effective et de viabilité des modèles économiques associés.

Comme cet ouvrage le montrera, le post-Snowden est loin d'être la première occasion où le projet de la (re-)décentralisation bénéficie d'un regain d'intérêt. Cependant, les questions de surveillance le présentent sous des nouvelles facettes, et relie très fortement la question de la décentralisation des réseaux à celle de la *privacy by design*, ou « protection de la vie privée dès la conception » : principe techno-juridique selon lequel toute technologie exploitant des données personnelles doit intégrer la protection de la vie privée dès sa conception, et s'y conformer tout au long de son cycle de vie. Il s'agit là d'un concept appelé à connaître un fort développement, notamment en Europe, dans le cadre de l'élaboration, actuellement en cours, du Règlement sur la protection des données.

L'Europe et la *privacy* : pour une protection inscrite dans les architectures ?

La protection des données personnelles « par la technique » constitue donc un enjeu qui fait actuellement l'objet de négociations-clefs au niveau européen. Si l'affaire Snowden a assuré une visibilité renouvelée à cet enjeu, l'histoire de ce dernier est bien plus longue ; notamment, l'argument que les Européens « ne sont pas maîtres de leurs données » revient régulièrement dans les médias depuis la création de la Directive européenne (95/46/EC) sur la protection des données personnelles en 1995.

Des discussions aussi sensibles que décisives sont en cours entre l'Union européenne et les États-Unis sur le « *safe harbour* », l'accord qui régit les transferts de données personnelles vers les entreprises américaines qui y adhèrent, ainsi que sur le « *umbrella agreement* » qui concerne les droits des citoyens européens lors de la transmission de données dans le cadre de procédures judiciaires. La question de l'échange des données se pose également dans le cadre du Partenariat transatlantique de commerce et d'investissement (TTIP), l'accord commercial en cours de négociation entre l'Union européenne et les États-Unis, qui vise à éliminer les barrières commerciales dans de nombreux secteurs économiques afin de faciliter l'achat et la vente de biens et de services entre l'UE et les États-Unis³.

³ http://ec.europa.eu/trade/policy/in-focus/ttip/about-ttip/index_fr.htm

Ce contexte est en arrière-plan des positions prises par le Parlement européen dans l'affaire Snowden, après neuf mois d'enquête : le Parlement a demandé l'accélération de la négociation sur la protection des données, la suspension de la décision *safe harbour* ainsi que de l'accord TFTP (échange de données financières en matière de lutte contre le terrorisme). La résolution du Parlement en relation avec ces enjeux appelle, notamment, à la création d'un « *habeas corpus* numérique »⁴.

Autant d'arènes où les alternatives décentralisées aux services Internet, en tant que moyen de définir et assurer de manières différentes la protection des données personnelles, pourraient être prises en considération par l'Europe. Par ailleurs, la grande occasion de leur porter un soutien actif est sans doute représentée par les négociations actuellement en cours sur le projet de Règlement de protection des données en Europe⁵. Ce texte, qui devrait remplacer la Directive sur la protection des données 95/46/EC, est explicitement reconnu par l'Europe comme une nécessité face aux « défis de la mondialisation et des nouvelles technologies »⁶.

Etant donné que des documents tels que le Règlement fixeront le cadre des législations nationales pour les années à venir, il semble important que leur traitement des questions de *privacy by design* inclue également des références explicites aux « nains » de l'Internet, aussi bien qu'à ses « géants ». Il convient de noter qu'actuellement, les dispositifs basés sur des architectures distribuées cherchent à s'insérer et à trouver leur place dans un scénario de « concentration » dominant qui risque à tout moment, par le droit, par la technique ou un mélange des deux, d'entraver des possibles solutions alternatives à la protection des données personnelles. Les architectures P2P pourraient-elles, dans ce cadre, constituer une voie stratégique originale à suivre en Europe en termes de protection des données personnelles ?

La pluralité des « modèles Internet » dans le débat public et politique ?

S'il semble difficile, de par leur nature distribuée et « discrète », d'inscrire la promotion de ces alternatives dans la loi, il semble non seulement possible mais opportun que la politique d'innovation européenne les soutienne activement, et pour ce faire, d'autres voies existent.

Il serait, tout d'abord, souhaitable que l'Europe soit prête à décourager de manière plus ou moins contraignante des projets de loi nationaux qui, comme il a déjà été

⁴ <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0230+0+DOC+XML+V0//FR&language=FR>

⁵ http://ec.europa.eu/justice/data-protection/index_fr.htm

⁶ *Ibid.*

le cas dans le passé, viseraient à entraver des usages spécifiques du P2P, posant une sérieuse menace à l'innovation dans ce domaine. Plus généralement, en sortant du cadre du P2P, il faudrait être prudent par rapport à toute proposition de loi visant à préserver des monopoles existants face à des alternatives basées sur les nouveaux usages numériques.

L'Europe a ensuite l'opportunité de développer ultérieurement un programme de recherche qui permette de faire avancer l'état des connaissances sur l'économie du partage, la production entre pairs, les architectures décentralisées⁷. Si les chercheurs peuvent être mobilisés dans les arènes politiques pour rendre compte des enjeux du distribué, il s'agit aussi de produire des prototypes, qui prennent en compte les dimensions sociales et organisationnelles en plus de la dimension technique. Le volet de ces recherches qui concerne la viabilité des modèles économiques pour les architectures décentralisées, qui peinent encore à trouver leur voie, devrait être tout particulièrement valorisé, car à l'heure actuelle, les solutions centralisées qui sont en place bénéficient d'une avance appréciable tant au niveau de la qualité de service, que de la masse d'utilisateurs, créant des effets d'inertie et de difficulté de «migration» particulièrement redoutables pour des solutions alternatives.

Il s'agit également, pour l'Europe, d'être capable d'envisager des stratégies «vraiment» alternatives. Face à la domination des *clouds* centralisés, dont les fermes de serveurs sont très souvent hébergées sur le territoire américain, il serait souhaitable que l'«indépendance technologique» de l'Europe, envisagée comme partie intégrante de son *habeas corpus* numérique⁸, comporte des solutions P2P et décentralisées, plutôt que de considérer comme alternative la promotion... du *cloud* centralisé européen.

C'est pourtant ce qui se passe à l'heure actuelle : dans le cadre des solutions traditionnelles, qui prévoient l'hébergement des données personnelles sur des serveurs appartenant aux entreprises, la question centrale pour la protection des données est de savoir où est localisé le serveur. Dans ce scénario, promouvoir une «alternative européenne» consiste à contraindre les géants du net à localiser leurs serveurs là où ils collectent les données, ou alors à favoriser des hébergeurs européens – en faisant au passage l'hypothèse que les services de renseignement européens seraient plus respectueux de la vie privée que leurs homologues américains. Dans le cas de l'adoption de solutions P2P, les fermes de serveurs disparaîtraient ou se verraient fortement redimensionnées (comme on verra tout au long de cet ouvrage, les solutions de P2P «intégral» sont rares et présentent des complications) : la localisation des serveurs cesserait donc d'être le nœud central de la controverse.

⁷ En sont des exemples P2Pvalue (<http://www.p2pvalue.eu/>), CONFINE (<http://confine-project.eu/>), D-CENT (<http://dcentproject.eu/>), P2P-Next (<http://www.p2p-next.org/>).

⁸ Voir note 4 et (Hunt, 2014).

Il s'agit enfin de rendre partie intégrante du débat public européen la question des modèles organisationnels et architecturaux que sous-tendent actuellement nos services Internet; le public – et les pouvoirs publics – sont de plus en plus conscients que le partage, le regroupement, le stockage des données dans les populaires services Internet d'aujourd'hui répondent à un modèle où le trafic est redirigé vers un ensemble de machines, placées sous le contrôle direct et exclusif du fournisseur de service. Mais ils sont beaucoup moins conscients de l'existence de toute une galaxie de services décentralisés où les opérations de traitement de données ont lieu, en priorité ou en totalité, sur et entre les ordinateurs des utilisateurs, liés directement les uns aux autres. Il s'agit là d'un choix d'architecture qui contribue à façonner des définitions particulières de la protection de la vie privée de l'utilisateur, en modifiant les manières dont le contrôle sur les données informationnelles et la responsabilité de leur protection sont répartis sur les utilisateurs, sur les fournisseurs de service, sur les concepteurs qui l'ont développé.

Un soutien actif de l'Europe à de telles architectures pourrait également passer par une campagne de sensibilisation à la responsabilité, individuelle et collective, que représente un plus grand contrôle sur ses données dans un modèle distribué, et aux risques ou désavantages éventuels que celui-ci pourrait représenter. En effet, si ce modèle remet en question «par la technique» l'étendue et l'ampleur des droits acquis par le fournisseur du service sur les données personnelles de l'utilisateur par rapport au contrôle que celui-ci conserve sur ces données, il s'agit là d'un compromis. L'utilisateur voit sa *privacy* renforcée par la possibilité d'un contrôle augmenté sur ses données et les opérations du client P2P dont elles font l'objet, mais en même temps et pour les mêmes raisons, la responsabilité de ses actions se voit augmentée, tandis que le fournisseur renonce de son plein gré à une partie de son contrôle. La dimension collective de cette responsabilité se voit aussi accentuée, dans la mesure où l'infraction au comportement collaboratif a des conséquences collectives aussi bien qu'individuelles – qu'il s'agisse d'un stockage de contenus inappropriés ou de l'introduction d'informations non fiables ou de spam dans un index de recherche distribué.

Si les architectures P2P peuvent, à certains égards, ouvrir la voie à des améliorations de la protection de la vie privée et de la sécurité, leur défi principal est de se confronter aux protocoles et standards les plus utilisés, à l'ombre desquels elles vivent actuellement – ce qui rend beaucoup plus difficile, en amont, toute possibilité d'adoption et diffusion à large échelle, à cause de l'effet «le gagnant prend tout» qui caractérise le marché des services Internet. Les architectures P2P font face à nombre de défis, notamment la difficulté à trouver des modèles économiques viables pour les services qui emprunteraient la voie de la décentralisation, ou encore la difficulté à persuader les utilisateurs à abandonner la commodité du cloud Google ou Facebook pour des procédures plus compliquées: en effet,

l'auto-hébergement et la mise en sécurité d'un service Internet décentralisé nécessitent à ce jour de beaucoup plus de temps et d'investissement.

L'opportunité politique et sociale des dispositifs et services décentralisés n'aura cependant jamais été aussi évidente qu'avec l'affaire Snowden. Le vrai défi pour ces alternatives semble donc celui de saisir cette opportunité et de se rendre, aux yeux d'une masse importante d'utilisateurs, aussi rassurantes et faciles à utiliser que les «géants» du *cloud* centralisé. La transition vers un Internet qui repose de façon plus importante sur des technologies P2P et décentralisées pourrait maintenant séduire en tant que stratégie de contournement des tentatives d'appropriation des données personnelles – tant de la part d'acteurs malveillants que d'agences gouvernementales ou entreprises privées trop intrusives. Les technologies décentralisées ont mené pendant des années – à l'exception des applications de partage de fichiers – une existence discrète : les révélations Snowden les ont ramenées au premier plan et proposées en tant que possible instrument de protection de la vie privée – protection par la technique, par l'architecture, *by design*.

Francesca Musiani
Paris, septembre 2014

Introduction

*Peer-to-peer*⁹ : une technologie de réseau qui ne cesse de provoquer à la fois enthousiasme et inquiétude. Sa définition technique est relativement simple : il s'agit d'un modèle de réseau informatique structuré de manière décentralisée, afin que les communications ou les échanges qui y ont lieu se fassent entre nœuds dotés d'une responsabilité égale dans le système. Les participants au réseau mettent à disposition une partie de leurs équipements et ressources informatiques (capacité de calcul, espace de stockage, bande passante) ; accessibles de manière directe par les pairs, ces ressources partagées sont nécessaires au bon fonctionnement du service offert par le réseau. La dichotomie entre un serveur, fournisseur de ressources, et les clients demandeurs de ressources, caractéristique du modèle *client-serveur*, est remplacée par une situation où tous les pairs hébergent ou fournissent la ressource, et tous les pairs la demandent.

Pour un très grand nombre d'utilisateurs de l'Internet – depuis la rencontre entre le P2P et le grand public propulsée par le logiciel de partage de fichiers Napster, en 1999 – cette technologie est *de facto* synonyme de téléchargement (illégal) de contenus culturels, tandis que pour d'autres, elle représente l'utopie ultime du techno-égalitarisme et suggère un modèle organisationnel durable pour les sociétés de demain. Les objets, les démarches de conception et développement, les collectifs et les usages dont il sera question dans cet ouvrage ne peuvent pas complètement faire l'économie des visions normatives fortes qui se confrontent et s'affrontent autour du P2P, qu'il s'agisse de l'identification de cette technologie avec le piratage de contenus culturels et intellectuels, ou à l'extrême opposé, de son déploiement dans une arène démocratique, égalitaire et collaborative accessible à tous : les histoires de P2P que raconte cet ouvrage sont informées par ces visions, discours, narratives, et les informent en retour.

S'il ne veut certes pas négliger pas la puissance d'agir¹⁰ de ces visions normatives, ce travail – issu d'une thèse de doctorat soutenue à MINES ParisTech – ne souhaite cependant pas être une contribution ultérieure aux débats sur le droit d'auteur et sur la dialectique pillage/partage auxquels le P2P semble désormais être « naturellement » associé. Ce travail prend comme son point de départ la

⁹ Souvent traduit en français par « pair-à-pair » et plus souvent abrégé en P2P, abréviation qui sera utilisée par la suite.

¹⁰ C'est ainsi que Serge Proulx traduit une des notions fondatrices de la sociologie des sciences et des techniques, l'*agency*, en soulignant que cette traduction met en relief à la fois la dimension collective de l'agir et le « pouvoir des acteurs de formuler une critique effective et ainsi, de *faire dévier* le cours de la reproduction du monde social » [Proulx, 2009].

caractéristique de base du P2P en tant que modèle de réseau informatique: ce qui est, on l'a anticipé, le fait de permettre des échanges de données efficaces et directs entre des nœuds égaux. Égaux en termes de leur mise à disposition de ressources techniques à l'ensemble du système, et de la responsabilité qui leur est attribuée dans son fonctionnement.

Cet ouvrage s'intéresse donc au développement et à l'appropriation de services basés sur Internet¹¹ dont la conception intègre un choix de design particulier: le fait de déléguer la responsabilité et le contrôle de la gestion des données et des flux aux marges ou à la périphérie de ces systèmes en réseau. Les opérations nécessaires au bon fonctionnement des systèmes, et au fait qu'ils fournissent correctement les services auxquels ils sont destinés, dépendent techniquement des utilisateurs – appelés ici les «nains» du réseau: leurs terminaux, leurs ressources informatiques, mobilisées de manière agrégée pour servir un objectif commun.

Pour autant, cet ouvrage ne touche que de manière secondaire au type de service qui est le plus souvent associé à l'architecture P2P, le partage de fichiers. Il souhaite s'intéresser à la rencontre entre le choix de développer une architecture technique en P2P et des usages tels que la recherche d'information, le réseautage social, le visionnage de vidéos, le stockage de fichiers en ligne. Il s'agit d'usages qui nous sont très familiers dans notre pratique quotidienne d'internautes et d'utilisateurs de services en ligne, sous le nom de Google, Facebook, YouTube, Dropbox – les «géants» des technologies de l'information, basés sur une architecture de réseau client/serveur qui préconise une dichotomie clairement identifiable entre un serveur fournisseur de ressources, et des clients qui en sont demandeurs. Ce travail explore des dispositifs qui, tout en répondant à ces mêmes nécessités d'usage – recherche, réseautage, stockage – ont en commun un aspect d'architecture technique original par rapport à leurs célèbres contreparties centralisées: tous sont basés sur des technologies de réseau en P2P.

En suivant et en essayant de clarifier le «ballet entre programmeurs, logiciels et utilisateurs» [Abbate, 2012] qui construit la décentralisation dans les services Internet, cet ouvrage explore les implications socio-politiques de l'approche distribuée et décentralisée à l'architecture technique des services Internet, en faisant l'hypothèse qu'une telle approche aux «couches inférieures» de ces

¹¹ Beaucoup a été écrit sur le concept de «service» en sociologie des organisations, du travail et de la communication, ainsi qu'en socio-économie de l'innovation [Callon, Méadel & Rabeharisoa, 2002; Callon, Millo & Muniesa, 2007]. On adopte ici une définition basique de service Internet, celle de mise à disposition, au moyen de l'Internet, d'une capacité de communication, d'accès à l'information, de mise en relation ou de calcul.

systèmes questionne, ou peut questionner, les usages qu'ils servent, les dynamiques qui y ont lieu, les démarches qu'ils comportent¹².

L'ouvrage introduit d'abord la décentralisation des services Internet comme «problème» et discute les manières dont les sciences sociales peuvent en explorer les facettes invisibles [chapitre 1]. Il propose ensuite une réélaboration de l'histoire du P2P [chapitre 2], pour montrer comment, en cherchant les meilleures façons de contourner les problèmes spécifiques où l'architecture de l'Internet est durement mise à l'épreuve par les scénarios de «concentration» actuels, les développeurs se retournent vers l'Internet d'il y a quinze ou vingt ans : le premier Internet qui avait été créé en tant que système de communication entre machines de statut égal, partageant des ressources l'une avec l'autre. Le cœur de l'ouvrage [chapitres 3, 4 et 5] explore tour à tour : les contraintes, techniques et de marché, qui s'exercent sur un dispositif dont les concepteurs veulent introduire une version de P2P «intégral» dans la recherche d'information en ligne ; la construction, dans un service de stockage, du compromis entre une approche largement dépendante d'une plateforme de serveurs, et une approche d'architecture distribuée et décentralisée ; la problématisation politique et technique du *peer-to-peer* en tant que modèle alternatif pour les services audiovisuels par Internet, ainsi qu'à sa légitimation comme *valeur* dans le contexte européen. Le sixième et dernier chapitre tire les «fils rouges» de ce travail, en cherchant, au moyen des cas d'étude, des réponses à deux questions : Qu'est-ce que dessine une architecture de réseau décentralisée du point de vue de l'articulation des acteurs et des contenus, de la répartition de responsabilités, de l'organisation du marché et de la capacité à exercer du contrôle, des formes d'existence et des rôles d'entités telles que les nœuds du réseau, les usagers, les unités centrales ? Sous quelles conditions un réseau qui répartit la responsabilité de son fonctionnement à ses marges, et suivant un modèle non hiérarchisé ou hybride, peut-il se développer dans l'Internet d'aujourd'hui ?

¹² Cette approche doit beaucoup aux discussions et aux travaux menés en collaboration avec Alexandre Mallard et Cécile Méadel dans le cadre du projet ADAM (Architectures distribuées et applications multimédias), ainsi qu'à des travaux menés en collaboration avec Valérie Schafer et Hervé Le Crosnier [Musiani & Schafer, 2011 ; Schafer, Le Crosnier & Musiani, 2011].

De l'intérêt des « tuyaux » de l'Internet

«Le P2P, c'est des tuyaux. La plupart des gens ne s'intéressent pas aux tuyaux,» soulignait il y a quelques années Dan Bricklin – le développeur de la première feuille de calcul, *VisiCalc* – dans un des premiers ouvrages portant sur le potentiel «perturbateur» des technologies P2P [Bricklin, 2001 : 59]. La «plupart des gens» auxquels Bricklin se réfère dans cette citation est constituée bien sûr par les utilisateurs des premières et populaires applications de partage de fichiers en P2P qui, comme Napster, connaissaient leur heure de gloire à l'aube du XXI^e siècle.

Dan Bricklin a très probablement eu raison dans son évaluation des raisons du succès auprès du grand public des applications P2P pionnières, destinées à servir le partage massif de contenus numériques : celui-ci doit son succès à la capacité de ces systèmes à trouver et obtenir rapidement un contenu spécifique qui suscite l'intérêt de l'utilisateur, plutôt qu'à l'architecture en pair-à-pair qui les sous-tend. Pourtant, cette remarque joue un rôle de premier plan dans la définition du regard que cet ouvrage porte sur des objets variés, dont le point commun – et la spécificité par rapport à leurs homologues centralisés plus célèbres – est la logique P2P qui sous-tendait leur modèle technique. Il s'agit de s'intéresser aux « tuyaux », ou, comme Susan Leigh Star l'a très bien formulé, de « faire remonter à la surface le travail invisible » [Star, 1999 : 385] sous-tendant les pratiques, les usages et les échanges en réseau, pour rechercher dans le design et la conception des « couches inférieures » d'un réseau certaines des raisons à la base de sa réappropriation par les usagers, de ses formes d'organisation et de régulation.

ARCHITECTURES : LA DÉCOUVERTE DU « TRAVAIL INVISIBLE »

L'architecture d'un réseau ou d'une application est la structure technique qui le ou la sous-tend [van Schewick, 2010], conçue selon une «matrice de concepts» [Agre, 2003] : c'est-à-dire, son agencement logique et structurel, fait d'équipements de transmission, protocoles de communication, éléments d'infrastructure, et connectivité entre ses composantes ou nœuds. Barbara van Schewick souligne que des décennies de recherches sur les processus d'innovation ont permis de mieux comprendre comment les changements du droit, des normes, des prix influencent le contexte économique de l'innovation et les décisions des innovateurs, et vice-