

LA DONNÉE NUMÉRIQUE,
BIEN PUBLIC
OU INSTRUMENT DE PROFIT

75

Les discours sur la révolution industrielle du numérique sont passés de la métaphore sur les autoroutes de l'information à celle sur la société de la connaissance. À partir du milieu des années 2000, la notion de mise en données du monde (*datafication*) a pris le dessus sur ces dernières, traduisant une prise de conscience progressive de l'importance croissante des données dans cette nouvelle économie qui se met en place depuis le début de la décennie 1990.

« Données », « informations » et « connaissances » sont d'ailleurs des termes souvent utilisés l'un pour l'autre, alors même que ceux-ci recouvrent des catégories bien différentes tout en s'imbriquant. Cette confusion, trop souvent entretenue, est d'ordinaire augmentée par une autre métaphore qui rajoute à la confusion : celle d'un nouvel eldorado pétrolier que constitueraient les données. Le succès commercial et financier de certaines entreprises numériques, principalement américaines, attesterait de cette réalité. Les chiffres sont éloquentes : les profits de Google pour 2016 s'élèvent à 19,5 milliards de dollars, contre 10,6 milliards quatre ans plus tôt. Dans le même secteur, celui de la publicité en ligne, Facebook réalise pour sa part un résultat net de 10 milliards de dollars en 2016 pour un chiffre d'affaires de 27 milliards. Facebook a multiplié son résultat par près de deux cents entre 2012 et 2016. Ces deux réussites emblématiques feraient donc office de preuve irréfutable de l'existence de ce nouvel eldorado et de la capacité de certaines entreprises à capturer la valeur des données.

Cependant, si les données constituent bien une nouvelle matière première, leurs caractéristiques et leurs mécanismes de production et de

valorisation différent profondément de ceux du pétrole. Une grande partie des données est le produit de nos interactions quotidiennes dans l'espace numérique et avec le monde physique, de nos échanges et transactions. Dès lors, la nature des données a une importance fondamentale dans la nouvelle économie qu'elles fondent. Elle renvoie à des problèmes qui ne sont pas purement économiques mais aussi politiques, puisqu'elle interroge l'organisation et la répartition de la valeur produite à partir des données, notamment lorsque leurs mécanismes de production et de collecte sont des plus opaques ou qu'ils empêchent la création de valeur. Les régimes juridiques qui encadrent le traitement des données sont alors essentiels pour comprendre les mécanismes de création de valeur correspondants. Ainsi le mouvement d'ouverture des données publiques, l'*open data*, est-il une alternative juridique et politique nouvelle qui

76 fonde une autre valeur de la donnée.

Pour contribuer utilement au débat sur la valeur des données, il nous semble indispensable de revenir avec précision sur la notion même de donnée, afin d'en saisir la nature et de lever les trop nombreuses confusions qui polluent le débat actuel. Si d'aucuns s'inquiètent de l'appropriation privée de cette valeur, force est de constater qu'au-delà de ce constat une analyse des mécanismes de valorisation de la donnée n'est jamais réellement produite. Aussi voudrions-nous nous y employer pour éclairer les débats sur la possibilité d'une alternative à la privatisation des données. Autrement dit, l'*open data* et les « communs de données » sont-ils des alternatives crédibles aux mécanismes d'appropriation privée de la donnée ? Plusieurs conditions nous semblent encore manquer pour que les données puissent devenir un bien public.

DE LA DONNÉE

De nombreuses analyses sur le rôle nouveau des données dans l'économie sont souvent obérées par l'absence d'une définition précise de cet objet et de ce qui le distingue de l'information et de la connaissance. De plus, les caractéristiques de la donnée sont rarement discutées, encore moins analysées, alors même qu'elles en définissent souvent la valeur.

D'un point de vue économique, la donnée est un bien non rival, ce qui signifie que sa consommation par un individu ne la fait pas disparaître. Cette caractéristique essentielle rend caduque la métaphore pétrolière puisqu'il ne peut y avoir épuisement de la ressource. Le phénomène de numérisation du monde crée chaque jour davantage de données ; c'est précisément une des dimensions de l'expression « *Big Data* ».

Cette caractéristique explique également pourquoi certains envisagent la donnée comme un bien public. Cependant, un bien public pur est un bien non rival et non excluable, c'est-à-dire qu'il est difficile de faire payer l'accès à ce bien (on parle de « non-excluabilité »). Les données ne relèvent pas, par nature, d'un bien public car il existe de nombreux moyens d'en rendre l'accès payant ; c'est la base de nombreux modèles d'affaires d'entreprises¹. À cet égard, la gratuité apparente de multiples services numériques repose bien sur une monétisation des données par les modèles économiques bifaces très largement répandus dans l'économie du Net². Les données s'apparentent davantage à un bien de club, c'est-à-dire un type particulier de bien collectif dont la consommation respecte le principe de non-rivalité mais qui a la possibilité d'être rendu excluable. Au-delà de cette première approche, comment définir la donnée ?

Une donnée est la description élémentaire d'une réalité – par exemple, le fruit d'une observation (le nombre de voitures par jour qui passent sur une route) ou d'une mesure (la température d'une pièce). La donnée est souvent définie comme étant dépourvue de tout raisonnement, supposition, constatation, probabilité. Elle est souvent qualifiée de brute. Si cette idée d'objectivité de la donnée est très partagée, nombreux sont ceux qui considèrent que la notion de données brutes est un oxymore³. Toute production de données est un processus qui repose sur une intention, une subjectivité⁴. Les données scientifiques sont souvent considérées comme des données brutes par excellence, alors même qu'elles font l'objet, dans leur immense majorité, d'un important travail de nettoyage (*data scrubbing*). Il y a là encore une différence essentielle avec la matière première qu'est le pétrole : les données sont issues d'une intention et d'un procédé de production ; elles n'existent pas par elles-mêmes. Aussi est-il préférable de parler de données primaires plutôt que de données brutes. La production de la donnée, avant même que celle-ci ne soit traitée, nécessite ainsi des processus spécifiques, des ressources et une intention.

Les discours sur les données ignorent également une dimension

77

1. Philipp Max Hartmann, Mohamed Zaki, Niels Feldmann et Andy Neely, « Big Data for Big Business ? A Taxonomy of Data-driven Business Models used by Start-up Firms », CambridgeServiceAlliance.eng.cam.ac.uk, mars 2014.

2. David S. Evans et Richard Schmalensee, *Matchmakers: The New Economics of Multi-sided Platforms*, Boston (Mass.), Harvard Business Review Press, 2016.

3. Lisa Gitelman (dir.), *Raw Data Is an Oxymoron*, Cambridge (Mass.), MIT Press, 2013.

4. Samuel Goeta, *Instaurer des données, instaurer des publics. Une enquête sociologique dans les coulisses de l'open data*, thèse de sociologie, ENST Paris, 2016.

fondamentale de leur économie : leur qualité. Selon le cabinet Gartner, plus de la moitié des programmes de gestion de la relation client et des projets d'entrepôt de données (*data warehouse*) échouent du fait de la mauvaise qualité des données⁵. Une étude a identifié cent vingt-deux entreprises en Europe et aux États-Unis qui, ensemble, avaient perdu plus de 1 200 milliards de dollars en raison de problèmes liés au même écueil⁶. De nombreux travaux de recherche ont analysé ces problèmes et proposent des démarches et modèles afin d'y remédier. Tous soulignent la nécessité d'avoir recours à des processus organisationnels spécifiques. Encore une fois, la valeur de la donnée n'existe pas *per se* et requiert un certain type d'organisation pour en garantir la qualité.

78 Mais est-ce bien réellement à partir des données que la valeur est créée ? Il faut bien reconnaître que nombre de services numériques, comme les services de messagerie électronique, d'écoute de musique en ligne, les logiciels de traitement de photographies, les places de marché électroniques, etc., ne se basent pas directement sur les données mais plutôt sur les métadonnées. Les métadonnées décrivent les données ; elles peuvent faire l'objet d'une normalisation internationale ou être conçues par les utilisateurs eux-mêmes (ajout d'un descriptif à une photo, par exemple). Les services numériques offrent des fonctions de classement, de tri, d'organisation des données qui sans les métadonnées n'existeraient pas. Ainsi, un morceau de musique numérique n'est qu'un fichier difficile à manipuler par un utilisateur si on ne lui adjoint pas un fichier décrivant l'artiste, l'album, l'année, le style musical, etc. La valeur d'usage de ces services repose donc sur la combinaison de données et de métadonnées. Les données, seules, sont donc sans réelle valeur puisque inutilisables sans recours aux métadonnées.

Par ailleurs, de nombreuses analyses focalisent leur attention sur les données personnelles que les échanges numériques mobilisent⁷. Pour pertinente et fondamentale que soit cette préoccupation, on ne peut négliger le fait que les services numériques sont construits non pas directement sur ce type de données, mais à partir de données d'usage et de navigation, et de celles relatives à l'environnement du service.

5. Cité par Adrian Gregory, « Data Governance – Protecting and Unleashing the Value of your Customer Data Assets », *Journal of Direct, Data and Digital Marketing Practice*, vol. 12, n° 3, 2011, p. 230-248.

6. Larry P. English, *Information Quality Applied: Best Practices for Improving Business Information Processes, and Systems*, Indianapolis (Ind.), Wiley, 2009.

7. Cf. par exemple « L'économie des données personnelles : les enjeux d'un business éthique », Cigref.fr, octobre 2015.

Ce sont en majeure partie les traces des interactions avec un dispositif numérique qui sont à l'origine de la valeur du service⁸. Initialement, ces données sont produites pour des raisons techniques en vue d'assurer le bon fonctionnement du service (comme les *cookies* ou les *logs* sur les serveurs, les coordonnées GPS pour la géolocalisation d'un service). Ces données, secondaires par rapport aux données de transaction, n'ont pas de valeur directe pour l'utilisateur. Cependant, dès lors qu'elles font l'objet d'un traitement, elles prennent une réelle valeur. Par exemple, le nombre de fois qu'un morceau d'un même artiste a été écouté sur un service de musique en ligne, les lieux ou les moments d'écoute, les listes de morceaux élaborées par les usagers, constituent des données qui sont utilisées pour bâtir des services complémentaires, comme les services de recommandations personnalisées. Nombreux sont les services numériques dont la valeur repose essentiellement sur l'exploitation de ces traces ancillaires, sans valeur pour celui qui les génère.

79

C'est par un traitement qu'une donnée primaire devient une information. Ce traitement nécessite généralement l'agrégation de plusieurs données pour construire une information. L'ensemble des transactions d'un magasin forment ainsi un agrégat, le chiffre d'affaires journalier. Cette information est interprétable par l'individu qui l'élabore. Elle est comparable à d'autres informations équivalentes. Elle a une utilité pour les individus dans la mesure où elle aide à prendre des décisions. Par conséquent, l'information possède une valeur plus élevée que les données.

En outre, la transformation de données en une information requiert des connaissances que traduisent le traitement sur les informations et l'interprétation qu'en ont les individus. Ces connaissances sont précisément l'objet d'une attention particulière de la part des entreprises depuis le milieu de la décennie 1990, ce qui a donné lieu à l'émergence du management des connaissances⁹. Les données n'ont donc pas de valeur sans recours à des connaissances organisées à des fins de traitement.

À partir du milieu des années 2000, le développement sans précédent des volumes de données grâce à la numérisation, combiné aux nouvelles capacités de calcul et de stockage, a permis de franchir une nouvelle étape

8. Emmanuel Kessous, « L'économie de l'attention et le marketing des traces », texte présenté lors du colloque « Web social, communautés virtuelles et consommation », 79^e congrès de l'ACFAS (université Bishop's et université de Sherbrooke), 11 mai 2011.

9. Ikujiro Nonaka et Hirotaka Takeuchi, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, New York (N. Y.), Oxford University Press, 1995.

dans le développement de l'intelligence artificielle. En l'état, les méthodes d'apprentissage automatique (*machine learning*) et d'apprentissage profond (*deep learning*) reposent massivement sur les jeux de données mis à disposition pour « entraîner » des algorithmes. Dès lors, la valeur des données n'est plus simplement liée à leur existence et à leur qualité, mais à leur combinaison dans des algorithmes auto-apprenants. Si les données sont décisives dans la création de valeur, les capacités de traitement et les compétences en matière de programmation le sont tout autant.

La nature de la donnée n'en fait donc pas, par elle-même, un bien public. Les données ne sont pas intrinsèquement sources de valeur. La donnée primaire nécessite des investissements de forme (qualité des données, métadonnées) pour que sa valeur puisse en être extraite. Plus encore, la valeur des données s'exprime *via* un processus de traitement
80 de plus en plus sophistiqué, dans lequel la valeur ne provient pas tant des données que de leur circulation et de leur combinaison dans des algorithmes et des procédés de médiation.

DE LA DONNÉE À LA VALEUR

Là où nombre d'études laissent penser que la seule collecte de données est source de valeur et de profits, l'analyse des caractéristiques de la donnée met en évidence une réalité plus complexe : des investissements nombreux sont nécessaires pour lui donner une valeur économique. Si certains identifient ces derniers aux investissements visant à élaborer des infrastructures informatiques, les mécanismes de création de valeur reposant sur la donnée requièrent des investissements en termes d'organisation et de compétences sûrement plus décisifs que les investissements matériels, ainsi que de nombreux travaux de recherche le mettent en évidence¹⁰. En outre, la valeur créée à partir des données ne sera perçue par l'utilisateur qu'à condition d'être médiée par des interfaces qu'il peut s'approprier.

En effet, la qualité des données ne peut être obtenue et maintenue dans le temps que grâce à une implication des différents acteurs qui gèrent des données au sein de l'entreprise concernée. Cela exige d'organiser des processus de gestion des données par l'intermédiaire d'un pilotage

10. Cf. par exemple Jeanne G. Harris et Vijay Mehrotra, « Getting Value from Your Data Scientists », SloanReview.MIT.edu, 16 septembre 2014 ; Sam Ransbotham, David Kiron et Pamela Kirk Prentice, « Beyond the Hype: The Hard Work behind Analytics Success », *ibid.*, 8 mars 2016 ; *id.*, « Minding the Analytics Gap », *ibid.*, printemps 2015.

adéquat. Une véritable « gouvernance de la donnée » est donc nécessaire pour garantir la qualité des données; loin d'être une simple préoccupation technique, elle doit être rattachée à la gouvernance globale de l'entreprise. À ce premier niveau s'articule un niveau managérial qui définit les processus de gestion des données (propriétaires de la donnée, mécanismes de gestion et de contrôle), qui s'articule lui-même à un dernier niveau, celui de la gestion technique des données. Le pilotage de ces trois niveaux peut être confié à un responsable des données, souvent dénommé *chief data officer*. Cette démarche devient essentielle car la volumétrie des données ne cesse d'augmenter et que la nature des données se diversifie. La question de la qualité des données aurait dû conduire les entreprises à développer une telle gouvernance de la donnée, mais c'est encore trop rarement le cas.

Cependant, il s'agit davantage d'un prérequis que d'une réelle capacité organisationnelle. Dès lors, sans une organisation adaptée et une capacité d'apprentissage, tenter d'exploiter des données donne souvent lieu à un échec¹¹. En outre, du fait de la nature même des méthodes du *Big Data*, qui s'appuient sur des volumes très importants de données, voire sur la totalité des données disponibles sur un sujet, il est nécessaire d'articuler la réflexion sur les mécanismes de création de valeur et le modèle d'affaires à celle sur les capacités techniques et analytiques afin que la conception des offres et des produits intègre bien la collecte de données. Ce qui implique de s'interroger sur la nature des compétences à réunir et la manière de les intégrer dans l'entreprise. Plusieurs ressources et compétences sont indispensables pour bâtir une capacité stratégique permettant d'extraire la valeur des données et de développer des modèles d'affaires fondés sur la donnée.

Les compétences requises renvoient à différents niveaux au sein de l'entreprise :

- stratégique: capacité à définir des modèles intégrant la donnée comme ressource centrale du processus de création de valeur;
- organisationnel: capacité à organiser les ressources et compétences de façon efficace, particulièrement grâce à une gouvernance de la donnée;
- technique: capacité à identifier les plateformes et ressources techniques les plus efficaces;
- managérial: capacité à gérer les différents métiers et à articuler les compétences en vue de créer de la valeur.

11. Sunil Mithas, Maria R. Lee, Seth Earley, San Murugesan et Reza Djavanshir, « Leveraging Big Data and Business Analytics », *IT Professional*, vol. 15, n° 6, 2013, p. 18-20.

Les seules compétences informatiques ne peuvent donc, à elles seules, garantir la capacité à créer de la valeur à partir des données. Il est nécessaire d'y adjoindre des compétences analytiques. Avec les volumétries nouvelles et l'évolution de la nature des données, celles-ci ont considérablement évolué depuis le milieu des années 2000 avec l'émergence du Web 2.0, qui a pour principale caractéristique de permettre aux utilisateurs eux-mêmes de produire des contenus (textes, images, vidéos, etc.).

82 Plus encore que l'identification de ces compétences, c'est la capacité à les intégrer dans les processus de décision au sein d'équipes pluridisciplinaires qui constitue le réel enjeu sur le plan managérial. Ce dernier point met en évidence le fait que la question de l'articulation des ressources est certainement primordiale pour assurer leur efficacité. Le design organisationnel est alors l'élément clé. Des travaux de recherche montrent d'ailleurs que ce n'est pas la seule existence de ces compétences mais leur répartition au sein de l'entreprise qui est à l'origine de la performance¹².

Dès lors, le processus d'extraction de la valeur est bien loin de se résumer à un assemblage technique de ressources et de compétences informatiques et analytiques. C'est une réelle question stratégique, car élaborer le design organisationnel des ressources et des compétences est essentiel pour obtenir un réel avantage compétitif par la mise en œuvre d'un modèle d'affaires fondé sur la donnée.

Quel que soit le processus de création de valeur adopté, il s'agit toujours d'un service rendu à un utilisateur ou à un client. Celui-ci valorisera ce service s'il satisfait son besoin, ou s'il le dépasse en donnant lieu à une expérience d'utilisation exceptionnelle. Dès lors, la capacité à intégrer l'utilisateur dès la phase de conception est un enjeu majeur dans la création de service numérique. Une des méthodes auxquelles on peut avoir recours pour y parvenir est le *design thinking*, développé à Stanford dans les années 1980 par Rolf Faste sur la base des travaux de Robert McKim¹³. Contrairement à la méthode qui caractérise la pensée analytique, le *design thinking* a pour but de développer un ensemble d'espaces qui s'entrecroisent plutôt qu'un processus linéaire comportant un début et une fin, et se pense comme une démarche itérative au sein de laquelle le prototypage joue un rôle important dans la quête d'une solution adéquate. Mais d'autres méthodes viennent en complément, telles

12. Prasanna Tambe, « Big Data Investment, Skills, and Firm Value », *Management Science*, vol. 60, n° 6, 2014, p. 1452-1469; Lorin M. Hitt, Fujie Jin et Lynn Wu, « Data Analytics Skills and the Corporate Value of Social Media. Kelley School of Business Research Paper No. 16-61 », 18 août 2016, disponible sur Papers.ssrn.com.

13. *Experiences in Visual Thinking*, Monterey (Calif.), Brooks/Cole, 1972.

que l'ergonomie d'interfaces. Sans adopter une telle approche, la restitution de la valeur créée par les données a peu de chance d'être effective.

Force est de reconnaître que la nécessité d'élaborer une interface appropriable n'est pas toujours prise en compte par les entreprises, en particulier lorsqu'elles ne perçoivent pas qu'une certaine part de la valeur se joue à cette étape du processus de création. Aussi la médiation de la valeur issue des données constitue-t-elle une étape souvent décisive dans le succès d'un service numérique reposant sur l'exploitation intensive de données et par là même de la valeur que les données contribuent à créer.

DE L'ORGANISATION POUR GÉRER LA VALEUR ISSUE DES DONNÉES

L'analyse de la chaîne de valeur de la donnée met donc en évidence sa forte complexité. Et plus encore le fait que la capacité à créer de la valeur à partir des données ne réside pas simplement dans les données elles-mêmes mais dans les investissements de forme et de procédés, notamment dans l'agencement des ressources et des compétences nécessaires à son exploitation. Dès lors se pose la question des formes d'organisation les plus aptes à créer de la valeur à partir des données.

83

Dans la première phase de la « numérisation du monde », ce sont essentiellement des organisations privées qui ont su maîtriser ces éléments et dégager des profits considérables dans un laps de temps très court. La vitesse à laquelle se déploie la révolution numérique n'a guère été relevée par les États et les corps intermédiaires de la société. Le réveil tardif de la puissance publique l'amène à découvrir le rôle nouveau de la donnée dans l'économie et la société sans que celui-ci ait réellement fait l'objet d'une réflexion stratégique approfondie de la part des responsables politiques. La vision qui s'est imposée à ce jour a donné lieu à une politique d'ouverture des données publiques (*open data*), d'une part, et à une régulation de l'utilisation des données personnelles en Europe avec le « règlement général sur la protection des données »¹⁴, qui entrera en vigueur en mai 2018, d'autre part.

Parallèlement, dans la lignée des travaux d'Elinor Ostrom¹⁵, une autre vision politique émerge qui considère l'information publique comme un bien commun et qui s'articule à la notion d'*open data* dans la mesure

14. Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016.

15. *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge, Cambridge University Press, 1991.

où les données générées selon ce modèle doivent demeurer ouvertes (tout en étant placées sous des licences d'utilisation qui peuvent varier d'un projet à un autre). Ce mouvement promeut un mode d'organisation de ces données en commun. On parle dès lors de données ascendantes ou contributives¹⁶, comme cela est le cas de celles créées par OpenStreetMap.fr en matière de cartographie ou OpenFoodFacts.org pour l'alimentation.

L'*open data* constitue une vision politique et économique de la donnée produite par les administrations publiques au sens large : ministères, agences publiques, collectivités territoriales, établissements publics à caractère industriel et commercial, délégataires de service public. Ainsi, les articles et bientôt les données scientifiques issues des recherches financées par des fonds publics en font partie.

84 Que signifie le fait que des données publiques soient ouvertes ? Une donnée ouverte est une donnée numérique dont l'accès et l'usage sont laissés libres aux usagers. Elle est diffusée selon une méthode structurée et placée sous une licence ouverte garantissant son libre accès et la possibilité de sa réutilisation par tous, sans restriction technique, juridique ou financière.

L'*open data* poursuit un objectif de transparence de l'action publique et vise également à fournir aux citoyens des données structurées (comme la base des adresses, qui donne accès à l'ensemble des adresses postales existantes en France), des données d'usage, des données sur les élus, mais aussi des données en temps réel relatives aux services publics (par exemple, pour signaler la disponibilité des véhicules d'autopartage à Paris).

Il a également un objectif économique lorsqu'il s'agit de créer de nouveaux services et un objectif en termes d'innovation publique quand il est question de co-crée à partir des données publiques. Dans cette perspective, l'État et les collectivités territoriales fournissent une des composantes des services futurs que des sociétés privées ou publiques pourraient concevoir et commercialiser.

Si l'ouverture des données publiques est une initiative relativement récente (2011), et que la loi pour une République numérique d'octobre 2016 acte le principe selon lequel les données publiques sont ouvertes par défaut, il convient néanmoins de constater que cette ouverture n'a pas à ce jour transformé les pratiques concernant la conception ou l'évaluation des politiques publiques, ni débouché sur des innovations

16. Marie Cornu, Fabienne Orsi et Judith Rochfeld (dir.), *Dictionnaire des biens communs*, Paris, PUF, 2017, art. « Données contributives ou ascendantes ».

majeures en termes d'action publique. Seules des initiatives locales de type *hackathon*¹⁷ ont ici ou là donné lieu au développement de quelques applications, représentations ou analyses nouvelles.

Plus encore, si les données sont désormais accessibles, elles le sont pour des citoyens dotés des compétences adéquates, notamment en matière de traitement. Aussi, s'il ne s'agit nullement de contester le bienfait de l'ouverture des données publiques, la question de leur utilisation de manière démocratique reste entière tant les compétences numériques sont loin d'être distribuées équitablement au sein de la population. S'il paraît utopique d'imaginer que chaque citoyen puisse s'emparer des données publiques par lui-même, il est toutefois possible d'inventer des médiations nouvelles pour ces données, en faisant émerger les compétences requises dans la société civile par une politique volontariste de développement de la « littératie numérique »¹⁸, comme le préconise le Conseil national du numérique¹⁹.

85

Les politiques publiques devraient faire émerger un écosystème d'acteurs intermédiaires, pas nécessairement des start-up, mais des acteurs aptes à s'emparer des données, à organiser leur analyse et à les restituer aux citoyens. Les politiques de médiation numérique devraient également avoir un tel objectif, afin que la société civile puisse s'emparer des données publiques et construire des analyses, représentations et modes d'action adaptés à la société numérique. Cela est d'autant plus nécessaire que l'intelligence artificielle, qui repose massivement sur l'utilisation de jeux de données pour entraîner ses algorithmes, se développe rapidement.

C'est un écosystème complet, autour de la donnée et de sa réutilisation par les administrations et la société, qu'il faut faire émerger pour instaurer un véritable État-plateforme, à l'instar des plateformes numériques qui bénéficient d'une logique de co-construction de la valeur grâce à leur écosystème particulier. Sans une telle action publique, il est à craindre que la création de valeur ne reste l'apanage des entreprises privées.

17. Événement au cours duquel un groupe de développeurs volontaires internes ou externes à une entreprise se réunit pour faire de la programmation informatique sur plusieurs jours en vue d'apporter une réponse nouvelle à un problème rencontré par l'entreprise. Dans ce cadre, cette dernière fournit des jeux de données aux participants.

18. Ce concept est défini comme la maîtrise de savoirs, de capacités et d'attitudes propres au domaine des technologies numériques (ordinateurs, tablettes, smartphones); il recouvre trois dimensions: la compréhension, l'utilisation et la création à partir des différents outils numériques. Cf. Michael Hoechsmann et Helen DeWaard, « Définir la politique de littératie numérique et la pratique dans le paysage de l'éducation canadienne », HabiMedias.ca, 2015.

19. *Citoyens d'une société numérique. Accès, littératie, médiations, pouvoir d'agir: pour une nouvelle politique d'inclusion* (rapport), 2013.

R É S U M É

La « mise en données du monde » s'est accélérée avec sa numérisation, provoquant l'émergence d'acteurs privés aptes à capter la valeur de données. L'examen des mécanismes de création de valeur à partir des données met en évidence la complexité de tels processus, ainsi que les investissements de forme nécessaires pour en extraire la valeur. Dès lors se pose la question de savoir si la seule ouverture des données publiques est en mesure de faire des données un bien public.