

# Quand les mondes de données sont redistribués : Open Data, infrastructures de données et démocratie



Jonathan GRAY

Chargé de cours en Études critiques des infrastructures,  
Département des Humanités numériques, King's College, London

L'open data, défini comme un ensemble d'idées et de conventions qui transforment l'information en une ressource publique réutilisable, est promu pour des objectifs variés : améliorer la transparence des institutions publiques, créer des projets qui renforcent la démocratie, stimuler la croissance économique. Les infrastructures sociales et techniques qui soutiennent l'open data recomposent les « mondes de données » : de nouveaux collectifs sociaux se forment, de nouvelles pratiques créatrices de sens apparaissent. Des initiatives politiques transnationales voient le jour. Loin d'être une simple « libération » des données, cela ne va pas sans traduction, médiation, et de nouvelles pratiques sociales. Mais ce mouvement peut-il servir de base d'une délibération démocratique plus riche, ou est-il voué à institutionnaliser socialement diverses formes de bureaucratisation et de marchandisation ?

Le pouvoir apparent et le potentiel de transformation des nombres a depuis longtemps inspiré de l'émerveillement, de l'inquiétude, et des actions de diverses sortes, pour le meilleur et pour le pire. La donnée est imaginée et utilisée pour comprendre, piloter et remodeler le monde, que ce soit pour stimuler la croissance économique ou pour redistribuer les ressources ; pour exploiter les ressources naturelles ou pour conserver les écosystèmes ; pour promouvoir la santé publique et l'éducation ou pour réprimer le crime ou la dissidence ; pour transformer la terre en territoire et les personnes en citoyens, consommateurs, travailleurs, camarades ou suspects. Des comptes rendus de tels projets peuvent être trouvés dans une littérature de plus en plus fournie d'histoire et de sociologie de la quantification et des statistiques<sup>1</sup>.

Qu'arrive-t-il à ces pratiques sociales et à ces imaginaires de la quantification quand des technologies numériques facilement disponibles facilitent la création, l'analyse et la reproduction

1. Voir par exemple :

- Hacking I. (1985), Making People Up, in T. C. Heller, M. Sosna, & D. E. Wellbery (Eds.), *Reconstructing Individualism: Autonomy, Individuality and the Self in Western Thought* (pp. 222-236), Stanford, Stanford University Press.
- Porter T. M. (1986), *The Rise of Statistical Thinking, 1820-1900*, Princeton, Princeton University Press.
- Porter T. M. (1996), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton, Princeton University Press.
- Desrosières A. (2002), The Politics of Large Numbers: A History of Statistical Reasoning, (C. Naish, Trans.), Cambridge, Harvard University Press.
- Espeland W. N. & Stevens M. L. (2008), A Sociology of Quantification, *European Journal of Sociology / Archives Européennes de Sociologie*, 49(3), 401-436. <https://doi.org/10.1017/S0003975609000150>
- Rottenburg R., Merry S. E., Park S.-J., & Mugler J. (Eds.) (2015), *The World of Indicators: The Making of Governmental Knowledge through Quantification*, Cambridge, Cambridge University Press.
- Bruno I., Jany-Catrice F., & Touchelay B. (Eds.) (2016), *The Social Sciences of Quantification: From Politics of Large Numbers to Target-Driven Policies*, New York: Springer.

des données par différents publics ? Quelles sortes de changements, de dynamiques, de controverses, de visions et de programmes peuvent-elles être observées quand les données deviennent numériques ? Une réponse récente à ces questions peut être trouvée dans le phénomène de l'open data qui peut être compris comme ensemble d'idées et de conventions visant à transformer l'information en une ressource publique réutilisable.

## Les conventions de l'open data

Par les pratiques et initiatives de l'open data, les sous-produits de l'administration, de la gouvernance et des autres activités des institutions publiques peuvent être transformés en une ressource « brute » - décrite indifféremment comme « le nouvel or », « le nouveau pétrole », ou « le nouveau terroir ». Cet objectif est poursuivi à travers une série de conventions légales, techniques et sociales visant à rendre les données publiques de manière à catalyser diverses formes d'innovation au-delà du secteur public et à travers lui. Ces conventions tirent parti d'une constellation de cultures et de normes associées à l'open source, au logiciel libre, à la culture libre, au piratage citoyen, au journalisme de données, aux données reliées, à la méthode agile de développement informatique et aux communautés du web 2.0. Si l'on transpose l'analyse d'Howard Becker sur les conventions qui font tenir ensemble « les mondes de l'art »<sup>2</sup>, quelles sortes de « mondes des données » ces conventions d'open data soutiennent-elles ?

Certaines de ces conventions visent à rendre les données légalement et techniquement réutilisables. Ainsi, les *licences ouvertes*, les régimes légaux et les politiques de l'information visent à atténuer les effets du droit d'auteur et des droits de propriété sur les bases de données qui peuvent entraver la réutilisation des données publiques, en stipulant clairement qu'elles peuvent être légalement réutilisées sans redevance ni autorisation. Elles tirent parti de pratiques légales associées avec le logiciel libre ou open source, la culture du libre, les groupes d'accès ouvert et de science ouverte. Il existe de fortes normes pour des formats de fichiers documentés publiquement, accessibles, structurés, qui donnent la priorité à des formats de données lisibles par les machines tels que « CSV » (« comma-separated values »), « JSON » (« JavaScript Object Notation ») et « XLS » (« feuille de classeur Excel ») par rapport à des formats de sortie imprimée comme « PDF ».

The screenshot shows the 'data.gov.uk' website interface. The main heading is 'Spend over £25,000 in the Cabinet Office'. Below this, there is a description of the dataset, a 'RESOURCE' section with a 'Download' button, and a 'TABLES' section with a table of data. The table has columns for Department, Agency, Year, Licence Type, Agency Name, Location, Procurement Number, Amount, and Description.

Department	Agency	Year	Licence Type	Agency Name	Location	Procurement Number	Amount	Description
Cabinet Office	Cabinet Office	2015	Open Data Licence	Cabinet Office	London	100110000000000000000000	100000000000000000000000	...
Cabinet Office	Cabinet Office	2015	Open Data Licence	Cabinet Office	London	100110000000000000000000	100000000000000000000000	...
Cabinet Office	Cabinet Office	2015	Open Data Licence	Cabinet Office	London	100110000000000000000000	100000000000000000000000	...
Cabinet Office	Cabinet Office	2015	Open Data Licence	Cabinet Office	London	100110000000000000000000	100000000000000000000000	...

Figure 1 : Copie d'écran du portail de données « data.gov.uk » montrant un aperçu de données de dépenses provenant du cabinet du Premier Ministre du Royaume-Uni.<sup>3</sup>

2. Becker H. S. (1984), Art Worlds, Berkeley, University of California Press.  
 3. On remarque les métadonnées de la licence d'open data en dessous du titre du fichier, le bouton de téléchargement des données brutes sur la droite, et un aperçu des données structurées dans la moitié inférieure de la page.

D'autre part, des centaines de « portails de données » locaux, régionaux ou nationaux émanant d'états, de citoyens, d'organisations non gouvernementales et d'entreprises réunissent des données de différentes sources pour les rendre plus faciles à trouver et à réutiliser. Par exemple, la figure 1 montre une page du portail « data.gov.uk » consacrée aux données de dépenses provenant du cabinet du Premier Ministre du Royaume-Uni ; cette page montre la licence ouverte favorisant la réutilisation, ainsi qu'un aperçu des données tabulaires structurées qui peuvent être téléchargées.

De même, les « journées de programmation informatique collaborative » ou « hackathons » visent à promouvoir la réutilisation des données ouvertes ; les « bourses », les « concours », les « incubateurs » et les « labs » font la promotion de l'innovation et de la collaboration. Ces dispositifs ont permis de donner le jour à des centaines « d'applications », de « projets de données », de « fonctions interactives », de « prototypes », de « sites Internet » et de « produits et services numériques ». Ils utilisent ces données ouvertes à des fins diverses, soit en créant de nouvelles cartes, des visualisations des données, des récits à partir des données, ou des « enquêtes de données », soit en personnalisant, adaptant, racontant, filtrant et combinant l'information d'autres manières.

## Un concept malléable

Ces développements ont été promus par différentes visions de ce qui pourrait advenir en ouvrant les données officielles. Certains suggèrent que cela peut améliorer la « transparence » et la « responsabilité » des institutions publiques, par exemple en créant des projets qui montrent comment les fonds publics sont dépensés. D'autres disent que les données aideront à augmenter l'efficacité du secteur public et à réduire ses coûts, d'une part en enrôlant des « auditeurs en fauteuil » qui identifieront le gaspillage et, d'autre part, en permettant et en encourageant des acteurs non-étatiques à produire des sites Internet et des services qui sans cela seraient produits sur fonds publics. D'autres soutiennent que les données publiques peuvent être utilisées pour créer des sites et des projets qui renforcent « la démocratie » - par exemple en permettant à des citoyens de contacter des institutions ou des politiciens, ou de se coordonner autour de tâches civiques ou collectives. D'autres maintiennent que les données peuvent être utilisées par de nouvelles firmes, des entreprises de technologie et des start-ups pour « créer des emplois » et « stimuler la croissance économique ». L'open data peut ainsi être compris comme un concept malléable, qui est reconfiguré pour s'aligner avec différentes conceptions des institutions publiques, des marchés et de la vie sociale<sup>4</sup>.

Comment est-ce que l'open data formate les pratiques sociales de quantification ? Ici, nous pouvons regarder au-delà des « fichiers de données » que valorisent les défenseurs de l'open data, et nous concentrer sur les « infrastructures de données » à travers lesquelles ces fichiers sont créés. Ces dernières sont des arrangements sociotechniques qui étayent la production des données : elles consistent en des écologies relationnelles de composants logiciels, des normes de données, des méthodes, des techniques, des comités, des chercheurs, des instruments et d'autres choses<sup>5</sup>. Plus généralement encore, les infrastructures de données créent des « mondes de données » dont les fichiers font partie dans lesquels différents acteurs voient les choses, s'en occupent et de s'y relient de différentes manières.

L'open data permet, et promet, différentes formes de « redistribution » et de « reconfiguration »

- 
4. Gray J. (2014), Towards a Genealogy of Open Data, Presented at the European Consortium for Political Research (ECPR) General Conference 2014, University of Glasgow. <http://dx.doi.org/10.2139/ssrn.2605828>
  5. Bowker G. C. & Star S. L. (1998), Building Information Infrastructures for Social Worlds — The Role of Classifications and Standards, in T. Ishida (Ed.), Community Computing and Support Systems (pp. 231–248), Springer Berlin Netherlands. Heidelberg.  
Bowker G. C. & Star S. L. (2000), Sorting Things Out: Classification and Its Consequences, Cambridge, MIT Press.  
Bowker G. C., Baker K., Millerand F., & Ribes D. (2009), Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment, in J. Hunsinger, L. Klastrop, & M. Allen (Eds.), International Handbook of Internet Research (pp. 97–117), Springer

de ces infrastructures de données et de ces mondes de données. Cela va des visions social-démocrates de participation populaire à des institutions publiques et aux services publics jusqu'à des politiques qui cherchent à limiter le rôle de l'état en lui faisant offrir ses données « brutes » que des acteurs non étatiques puissent utiliser comme base de leurs propres produits et services<sup>6</sup>. En tous cas, la distribution numérique des données signifie que le nombre des personnes pouvant y accéder et les utiliser est, en principe, démultiplié afin d'inclure quiconque dispose d'une connexion Internet et des connaissances de base. Les bases de données et le stockage de données, les technologies d'analyse et de visualisation montrent que les contextes d'usage des données du secteur public peuvent s'étendre bien au-delà des statisticiens, administrateurs, gestionnaires, chercheurs et fonctionnaires qui sont impliqués dans leur production et dans leur utilisation à l'intérieur des institutions publiques. De nouveaux « styles de raisonnement » (comme le dit Hacking) et de nouvelles manières de produire du sens sont rendus possibles du fait que les données transitent depuis les rapports administratifs vers des applications mobiles et des graphiques interactifs en ligne.

Au fur et à mesure que les fichiers sont recombinaés, extraits, appariés, reformatés, reconstruits, réutilisés et que différents acteurs leur donnent des sens différents, de nouveaux mondes des données émergent. La redistribution de ces mondes de données peut être comprise d'au moins trois manières<sup>7</sup>.

## De nouveaux collectifs sociaux

Premièrement, nous pourrions considérer la redistribution des mondes de données en termes de changement de composition des « mondes sociaux » de « expérience des données » et de « travail sur les données » comme une réalisation distribuée, collective. Les initiatives d'open data aspirent à provoquer des redistributions dans les collectifs associés à l'information publique, en s'adressant explicitement à de nouveaux acteurs au-delà du secteur public (qu'il s'agisse de citoyens, de groupes de la société civile, d'étudiants, de chercheurs, de journalistes, de start-ups ou d'entreprises de technologie) à travers des mécanismes comme les réseaux sociaux, les rencontres sur le web, les listes de diffusion, les « labs », les incubateurs, les hackathons, les programmes de bourses, les initiatives d'appel à la foule, les applications, les projets de données, les « sprints » de données, les « expéditions » de données, et les sites dédiés. Les portails de données invitent les utilisateurs à télécharger les données publiques et à réaliser leurs propres applications, visualisations, sites Internet et services. Les initiatives d'open data visent à exploiter la sagesse de la foule, l'expertise et l'expérience de différents acteurs en dehors de l'état, de façon à, comme l'a exprimé une initiative du gouvernement du Royaume-Uni, « Montrez nous une meilleure voie ». <sup>8</sup> Ainsi, nous pourrions examiner à la fois « les pratiques » et « les imaginaires » de la participation publique, de l'engagement public, de la co-création et de l'innovation autour des données publiques. De même, notre attention peut se porter sur les effets de ces mutations, des nouvelles configurations de politique économique jusqu'à des formes émergentes d'espace public, de mobilisation, de controverse sociétale et de formation de problèmes. Nous pouvons encore examiner les « imaginaires de données » et le « langage des données ». Pour finir, nous pouvons regarder quelles sortes de collectifs sociaux – les « publics de données » – sont effectivement assemblés autour des données publiques en pratique, ainsi que les politiques et le modelage des dispositifs et des procédures de l'engagement public, et le degré d'inclusion et d'exclusion de ces publics.

6. Gray J. (2014) cité

7. Gray J. (2018), Three Aspects of Data Worlds, *Krisis: Journal for Contemporary Philosophy*.

8. Voir : <http://news.bbc.co.uk/1/hi/technology/7484131.stm>

## Des pratiques créatrices de sens

Deuxièmement : nous pouvons considérer la recomposition des mondes de données en termes de changement dans « la manière dont les choses sont rendues compréhensibles », au sens où elles fournissent les conditions de possibilité de l'expérience du monde permettant de le comprendre, d'interagir avec lui et de lui donner sens. Pour paraphraser Bruno Latour : changez les instruments, et vous changez la manière dont les choses sont visibles, conservables et faisables avec les données. Dans ce cas, nous pouvons regarder les nouvelles pratiques créatrices de sens associées à l'open data, principalement l'agrégation et la combinaison de données de différentes sources. On s'intéressera alors à l'usage du « machine learning », des algorithmes et des nouvelles techniques analytiques ; ou encore aux nouvelles sortes de visualisations des données et « d'expériences des données » ; ou, pour finir, à des produits et dispositifs technologiques qui facilitent les différents modes d'entrée en relation avec les données – tels que les plateformes en ligne, les applications mobiles, la géolocalisation, le balisage, l'annotation, la production participative, les notifications en temps réel, la réalité augmentée et virtuelle, les technologies portables comme des vêtements, et les installations multimédia immersives.

Cette perspective met en évidence le fait que l'open data ne change pas seulement qui peut « utiliser » différentes ressources de représentation au sujet du monde, mais il facilite aussi de nouvelles sortes de « pratiques créatrices » pour donner sens à la vie collective. Par exemple, l'application « Walkonomics » (figure 2) combine de multiples sources de données publiques ouvertes avec des données générées par l'utilisateur pour fournir des évaluations de la « marchabilité » de différentes régions et de différents itinéraires dans plusieurs grandes villes – fournissant un nouvel outil pour collectivement quantifier, classer les lieux urbains et s'y relier.



**Figure 2 :** L'application « Walkonomics » pour iPhone et Android suggère la « marchabilité » de différents itinéraires en combinant différentes sources de données<sup>9</sup>

9. Voir : <http://www.walkonomics.com/> et <https://data.gov.uk/apps/walkonomics-find-walkable-route>



## Des projets transnationaux

Troisièmement : nous pouvons regarder comment l'open data remodèle les « projets politiques qui font le monde » (et peut-être plus récemment des projets qui « défont » le monde), comme ceux que mettent en évidence les recherches récentes sur les circuits transnationaux de la globalisation. Les projets d'open data transnationaux facilitent les nouveaux régimes de la quantification transnationale et l'agrégation, l'harmonisation et la standardisation des données.

Par exemple, « Open street map » permet aux utilisateurs d'ajouter des données géospatiales provenant d'agences publiques à des données qu'ils ont collectées eux-mêmes ou tirées de cartes<sup>10</sup>. De même, Le « Partenariat ouvert des marchés » crée de nouvelles normes pour que les données du secteur public soient « partageables, réutilisables, lisibles par les machines » et soient donc harmonisées et comparables à travers les frontières.<sup>11</sup> Autres exemples, citons le projet « Open spending » qui agrège des millions de transactions de dépenses depuis plus de 70 pays<sup>12</sup> ; ou encore « Open ownership » qui vise à créer un nouveau répertoire global de « qui contrôle les entreprises et qui en tire profit ». <sup>13</sup> Ces exemples peuvent être interprétés comme de nouveaux réseaux transnationaux d'expertise, d'échange et de transfert de connaissances permettant d'harmoniser et de standardiser le « travail des données » par-delà les frontières.

Pour conclure, j'espère que ces trois façons d'examiner la redistribution des mondes de données – en termes de collectifs sociaux, de pratiques créant du sens et de réseaux transnationaux – permettront d'attirer l'attention sur les différents aspects de la politique d'open data et d'information publique, et sur la manière dont les technologies numériques donnent lieu à différentes pratiques sociales et différents styles de quantification. Alors que la rhétorique de l'open data met souvent l'accent sur « la libération » de données préexistantes du secteur public pour alimenter l'innovation et l'extraction de valeur, nous espérons avoir montré que cela ne va pas sans traduction, médiation, et de nouvelles sortes de pratiques sociales et de mondes sociaux. Le passage des données du secteur public à divers autres acteurs donne naissance à de nouvelles sortes de « mondes de données ».

## D'autres mondes possibles ?

Un mouvement qui a commencé sa vie en appelant à l'ouverture des « fichiers » officiels peut encore contenir les germes d'un programme politique plus ambitieux visant à ouvrir « l'espace public, l'imagination, la participation, la délibération, la contestation et la créativité » autour de la fabrication des données<sup>14</sup>. Ce qui pourrait être relu en termes d'une « institutionnalisation » populaire de certains styles de raisonnement, modes d'épreuve et genres de problèmes de quantification (hérités comme un sous-produit de l'administration et de la gouvernance du secteur public) pourrait cependant servir de base d'une délibération démocratique plus riche et significative autour de « quelles choses » sont prises en compte à travers les données, « comment », et avec quels effets. En un temps où, dit-on, la confiance du public dans les institutions décroît, et où pourtant l'échelle des grands problèmes collectifs auxquels nous faisons face est considérée comme sans précédent dans les temps modernes, de telles expérimentations publiques autour du rôle des données dans les sociétés démocratiques doivent sûrement être bien accueillies. Si l'étude de l'open data suggère sans aucun doute les façons dont ce développement peut être utilisé comme un moyen d'accélérer et d'institutionnaliser socialement différentes formes de marchandisation et de bureaucratisation, elle peut aussi, parfois, nous récompenser avec des rappels que d'autres mondes des données sont possibles.

10. Voir : <https://www.openstreetmap.org/>

11. Voir : <https://www.open-contracting.org/about/>

12. Voir : <https://openspending.org/>

13. Voir : <http://openownership.org/>

14. Gray J. (2016), Datafication and Democracy: Recalibrating Digital Information Systems to Address Societal Interests, *Juncture*, 23(3).  
Téléchargeable depuis <http://www.ippr.org/juncture/datafication-and-democracy>