

Chapitre 1

Les racines multiples de l'open data

Le mouvement que l'on désigne aujourd'hui sous le terme d'open data est ancré dans des concepts, des pratiques et un cadre réglementaire qui anticipent largement son avènement. Avant de retracer l'émergence des principes de l'ouverture des données, essayons de comprendre les différentes filiations qui ont précédé l'apparition de cette notion. Les injonctions à l'ouverture des données se sont réclamées d'une multitude de mouvements et de principes dont la mise en évidence permettra de dépasser l'apparente uniformité des acteurs qui se sont réunis autour du terme open data.

Dans ce chapitre, nous aborderons cinq facettes de l'ouverture des données. Premièrement, la demande de transparence, qui apparaît aujourd'hui comme une revendication essentielle en démocratie, a donné lieu à des lois de droit d'accès à l'information publique qui accordent aux citoyens un accès direct aux documents administratifs. Le droit d'accès constitue le socle juridique de l'ouverture des données mais il porte sur l'ensemble des documents administratifs (dont font partie les données) et repose sur des demandes là où les politiques d'open data impliquent une approche proactive dans laquelle l'administration met à disposition volontairement les données sans attendre une requête. Deuxièmement, l'ouverture des données s'inscrit dans la lignée de la pensée cybernétique pour qui la libre circulation de l'information constitue une condition essentielle de la survie et de l'équilibre des sociétés démocratiques. Cette croyance dans le partage de l'information et du savoir a fait émerger le mouvement du logiciel libre dont les principes ont constitué la matrice de ceux de l'ouverture des données. Dans cette lignée, une multitude de mouvements se sont réclamés de la philosophie de l'ouverture pour réclamer la liberté de distribuer et de modifier des œuvres de l'esprit. Les données, en tant que fondement de l'information et du savoir, constituent pour ces acteurs un objet essentiel pour renouveler notre connaissance du monde. Troisièmement, l'ouverture des données tire une de ses sources dans les pratiques anciennes de partage de données en science. Dans certaines disciplines comme l'astronomie, la botanique ou la génétique, le partage des données et la réutilisation de données produites par d'autres constitue une pratique scientifique légitime et courante. C'est dans ce contexte que le terme « open data » a connu ses

premières utilisations pour désigner des pratiques de partage de données scientifiques. Quatrièmement, les industries de l'information ont réclamé de longue date la mise à disposition gratuite et la libre réutilisation des données du secteur public. Aux États-Unis, depuis la fin des années 1980, suite à une campagne de lobbying, les données produites par les acteurs publics sont librement réutilisables. En France, le Minitel a dégagé des revenus importants pour les ministères à partir des données publiques et a suscité des débats sur l'opportunité de maintenir des redevances. Espérant trouver un gisement de la croisement de croissance, la Commission Européenne a joué un rôle important dans l'arrêt progressif de la tarification des données. Enfin, l'ouverture des données a été souvent mobilisée comme un levier pour transformer l'administration et obtenir sa « modernisation » en cassant les silos organisationnels et en favorisant la collaboration avec la société civile. L'ouverture des données peut donc être prise avec des points de vue très différents, examinons là sous toutes ses facettes.

La transparence et le droit d'accès à l'information

Selon Clare Birchall {Birchall:2012hz}, la demande de transparence, entendue comme divulgation de l'information au public, a été popularisée principalement au XVIIIe siècle par des penseurs comme Kant, Rousseau ou Bentham. La transparence administrative, permise par le droit d'accès à l'information publique, trouve en Suède son acte de naissance. En 1766, le Parlement de Suède a adopté un décret relatif à la liberté de la presse et au droit d'accès aux documents publics qui prévoit que tout citoyen suédois est autorisé à consulter les actes officiels, un document étant "officiel" dès lors qu'il est conservé dans une administration {Ader:2005vp}. En France, avec la Déclaration des Droits de l'Homme et du Citoyen (DDHC) de 1789, ce droit d'accès a été revendiqué comme un élément indispensable de l'exercice de la souveraineté démocratique. L'article 15 de la DDHC qui stipule que « la Société a le droit de demander compte à tout agent public de son administration » pose les bases d'un devoir de réédition des comptes de l'administration aux citoyens. Ben Kafka, dans *The Demon of Writing* {Kafka:2012vt}, attribue l'origine de cet article à l'Abbé Sieyès. Homme d'église et essayiste, il a été un des principaux acteurs de la Révolution Française et l'auteur de plusieurs pamphlets dans lesquels il a esquissé les contours d'une société démocratique. Selon Sieyès, la démocratie représentative ne peut s'exercer que si les représentants ne rendent des comptes aux citoyens qui ont délégué l'exercice du pouvoir. Pour s'assurer que la souveraineté nationale soit bien exécutée, il faut selon lui comprendre dans quelles conditions les représentants ont exercé le pouvoir qui leur a été délégué. Ben Kafka

montre que le choix des termes de l'article 15 de la DDHC est beaucoup plus ambitieux que la constitution américaine dont l'article 1 réduit la réédition des comptes aux aspects budgétaires. Birchall souligne aussi que l'émergence de la notion de transparence doit beaucoup au développement au XIXe siècle des archives, bibliothèques et musées qui ont occupé progressivement une place centrale dans les villes et ont rendu les faits accessibles au public. Du fait de leur importance croissante au XXe siècle, les Etats-Unis ont fortement contribué à la popularité de l'idée de transparence. Le juge américain Louis Brandeis, a écrit dans son essai *Other People's Money* {Brandeis:KcT6--1i} cette formule restée célèbre au sujet de la transparence : « la lumière du soleil est dit-on le meilleur des désinfectants ; l'éclairage électrique le plus efficace des policiers. » Critiquant les banquiers de son époque accusés de conflits d'intérêt dans leurs investissements, Brandeis plaidait en faveur de la "publicité", l'idée que les banquiers doivent porter à la connaissance du public les commissions ou profits qu'ils perçoivent {Lessig:2009tq}. Symbolisant l'importance de la circulation des données dans l'intégrité publique et la lutte contre la corruption, cette phrase est aujourd'hui devenue le slogan de nombreux militants de l'open data dont la Sunlight Foundation (qui en tire son nom) ou Joshua Tauberer {Tauberer:2014td}. Adhérant aux idées de Brandeis, le président Wilson a nommé le juge à la Cour suprême et a placé la transparence de la diplomatie comme une condition essentielle de la paix dans ses « quatorze points », son programme pour mettre fin à la Première Guerre mondiale en 1918.

Jusqu'au milieu du XXe siècle, la transparence était un fondement théorique des démocraties mais reposait sur le truchement d'institutions comme les archives ou la Cour des comptes qui assurait pour le compte des citoyens la diffusion des informations publiques. Aux États-Unis, à la suite de la Seconde Guerre mondiale, le mouvement de l'*Open Government* a obtenu l'instauration d'un droit d'accès direct des citoyens aux documents et données produits par les acteurs publics {Yu:2012wv}. Portée en grande partie par des journalistes, l'ouverture des gouvernements désigne ici la révélation par l'État des secrets de son fonctionnement. Cette revendication est née de la critique de l'opacité de l'armée américaine lors de la Guerre froide et particulièrement pendant la Guerre du Vietnam. Autour de la notion de gouvernement ouvert, les journalistes tentaient alors d'obtenir un droit d'accès à l'information publique, un « droit de savoir », qui est devenu progressivement la norme dans la plupart des démocraties occidentales. En 1966, le *Freedom of Information Act* signé par le président Johnson a donné à chaque citoyen américain le droit d'exiger à l'administration les informations publiques qu'elle détient. En France, la loi CADA du nom de la Commission

d'Accès aux Documents Administratifs établit en 1978 un droit similaire qui permet à chaque citoyen d'exiger à une administration la publication des informations publiques. L'exercice de ce droit s'effectue par le dépôt d'une requête officielle auprès d'un agent public, la commission pouvant arbitrer en cas de litige. Il est important de préciser que les données qui sont sujettes à ces dispositions sont des informations publiques, produites dans le cadre d'une mission de service public. Le droit d'accès en France protège un certain nombre de secrets dont les délibérations du gouvernement ou les données à caractère personnel. Aujourd'hui, selon l'ONG Access Info, près d'une centaine de pays ont adopté une législation de droit d'accès à l'information.

Par rapport à ces dispositions, l'ouverture des données publiques s'inscrit directement dans la lignée du droit d'accès à l'information mais pose un principe de diffusion proactive. Là où le droit d'accès suppose une demande, l'ouverture consiste à publier de manière volontaire les données sans attendre une requête. Le mouvement de l'open data a aussi mis en lumière les données comme les éléments de base à partir desquels l'information et le savoir sont créés{Kitchin:2014ti. En diffusant les données dans leur plus grande précision, sous leur forme granulaire, les asymétries d'information entre le public et l'administration devraient se réduire et des connaissances inédites devraient émerger. La publication de données ouvertes est ainsi souvent considérée comme une forme supérieure de transparence, réputée plus objective que les formes narratives ou interprétatives de révélation {Birchall:2014eo}. Dispensées d'un biais narratif ou interprétatif, les données « brutes » sont associées à une objectivité qui configure les politiques d'open data comme le renouveau voire le paroxysme de la transparence de l'État {Goeta:2015tl}. Pour Evelyn Ruppert{Ruppert:2012ua}, la diffusion de données fait émerger des « publics des données » (*data publics*) qui sont institués par les politiques d'open data en inspecteurs du fonctionnement de l'État. S'intéressant au *Transparency Agenda* mis en place par le gouvernement Cameron en 2010 suite au scandale des dépenses des parlementaires, Ruppert a montré la dimension performative de l'ouverture des données : les politiques d'open data réalisent ce qu'elles énoncent. Dans l'esprit du Transparency Agenda, le simple fait d'exposer les données positionne les militants associatifs, les chercheurs ou les journalistes comme des « auditeurs de salon » (*armchair auditors*). Peu importe si ces différents publics s'empareront de ces données et sont en capacité de s'en saisir, la publication de ces données suffirait à atteindre un nouveau degré dans la transparence et le contrôle citoyen de l'action publique. L'ouverture des données marque ainsi une reconfiguration de l'expertise selon Ruppert. Alors que le Premier ministre David Cameron avait annoncé la naissance d'une nouvelle ère

post-bureaucratique où les agents publics ne sont plus les seuls à interpréter les données, elle met en garde sur le fait que « le pouvoir sur les données va ostensiblement reposer sur un public imaginé » alors que « la plupart des gens manquent des outils ou de l'expertise pour donner du sens aux téraoctets de données diffusés. »

La liberté de l'information et le logiciel libre

Au delà des revendications de transparence de la vie publique, la demande d'ouverture des données s'inscrit dans un mouvement plus large qui considère que la diffusion de l'information est impérative pour la survie des sociétés démocratiques. L'idée d'ouverture émerge au milieu du XXe siècle dans les écrits de philosophes et de théoriciens comme Karl Popper qui, dans *The Open Society and Its Enemies* publié en 1945, dénonçait les « sociétés fermées » fondées sur des vérités incontestables {Tkacz 2012; Russel 2014}. L'influence de la pensée cybernétique a aussi joué un rôle déterminant dans l'émergence de l'idéologie de l'ouverture. La théorie cybernétique postule que l'information est indispensable à la perpétuation de la civilisation. Dans un monde menacé par l'entropie, le chaos inéluctable annoncé par les principes physiques de la thermodynamique, les sociétés « ouvertes » parviennent à créer de l'ordre et à lutter contre l'effondrement {Breton:2004wa, Lafontaine:2004ws, Triclot:2008ul, Turner:2008ve}.

Cette croyance dans la circulation de l'information comme vecteur d'harmonie et de stabilité se retrouve dans le schéma de la pyramide du savoir (figure 1) qui place les données comme le fondement de l'information et du savoir {Rowley:2007gy, Fricke:2009ba, Floridi:2011tv}.



Figure 1. Schéma de la pyramide du savoir.

Influencé largement par la théorie cybernétique, Russell Ackoff, son auteur est un théoricien des organisations et du management de l'information. Il a proposé sa hiérarchie de l'information en 1989 dans la revue *Journal of Applied Systems*

Analysis {Ackoff:1989wy}. La forme de la pyramide indique une progression : les données seraient le fondement de l'information dont découle ensuite le savoir et même la sagesse. Elle indique aussi le poids relatif de chacune des strates comme l'expliquait son auteur dans une interview :

« Il y a cinq types de contenus dans l'esprit humain : les données, l'information, le savoir, la compréhension et la sagesse. Il y a une hiérarchie. Il y a un vieil adage académique qui dit qu'une once d'information vaut une livre de données, une once de savoir vaut une livre d'information, une once de compréhension vaut une livre de savoir, une once de sagesse vaut une livre de compréhension. Ça fait qu'une once de sagesse vaut environ 32 onces de données¹. »

Ce modèle est évidemment critiquable : la philosophie, dont l'objet consiste à atteindre une forme de sagesse, doit-elle être nécessairement fondée sur des données ? La collecte massive et la circulation des données sur un phénomène comme le réchauffement climatique a-t-elle entraîné mécaniquement plus de sagesse dans nos comportements ? Malgré ses limites, ce modèle a beaucoup circulé au point de devenir un des fondements des sciences de l'information.

Les mouvements liés au logiciel libre ont aussi largement milité pour la liberté de circulation de l'information. Sébastien Broca {Broca:2013uv} dans *L'Utopie du logiciel libre* revient sur les fondements intellectuels de ces mouvements et soutient que « la libre circulation de l'information est la valeur la plus communément associée au combat du Libre, que ce soit par les universitaires ayant étudié le sujet ou par les hackers eux-mêmes. » Des partisans du logiciel libre se sont en effet engagés pour que la connaissance et les œuvres bénéficient des mêmes libertés d'utilisation, de modification, de copie et de redistribution que le code informatique des logiciels qu'ils utilisent. Dans la continuité de luttes contre les verrous à la circulation de l'information, des libristes ont participé à la définition des principes de l'open data en défendant notamment l'usage de formats aux spécifications ouvertes et de licences dites « *copyleft* » qui empêchent légalement ce que la littérature autour des biens communs appelle l'« enclosure » des données.

Dans *Open Standards and the Digital Age*, Andrew Russel {Russel:2014um} invite à remonter plus loin que la filiation évidente du logiciel libre pour comprendre le foisonnement contemporain de mouvements qui se revendiquent de l'ouverture (*open access, open science, open government...*). Selon Russel, l'histoire des standards de télécommunications a fait émerger une philosophie voire une idéologie de l'ouverture. Dans les années 1970, des ingénieurs ont affirmé la

¹. Interview par Phyllis Haynes disponible en ligne : <https://www.youtube.com/watch?v=MzS5V5-0VsA>. Traduction personnelle.

supériorité des systèmes ouverts lors de la conception des standards mondiaux de télécommunication. En particulier, en août 1977, l'*International Organisation for Standardization* (ISO) a créé un groupe de travail intitulé *Open Systems Interconnection* (OSI) pour concevoir des standards ouverts d'interconnexion des réseaux et des terminaux. Bien que l'OSI a échoué dans sa mission, la métaphore de l'ouverture s'est imposée auprès d'ingénieurs issus de domaines très variés comme une critique de l'ordre établi et un cri de ralliement.

A travers ces multiples filiations (cybernétique, logiciel libre, standards ouverts), les militants de l'open data renouvellent la croyance en la circulation de l'information. L'ouverture des données pourrait rebattre les cartes et mettre sur un pied d'égalité l'ensemble des acteurs. En ayant accès aux données brutes, au matériau de l'information et du savoir, chacun serait en capacité de produire de nouvelles connaissances. Pour Kitchin (2014), « celui qui a accès à de larges données de grande qualité a un avantage compétitif sur ceux qui sont exclus de pouvoir générer de la compréhension et de la sagesse. » Si le savoir c'est le pouvoir², ouvrir les données serait un moyen efficace de le distribuer.

Le partage des données scientifiques

En plus de s'inscrire dans les revendications de transparence et de circulation de l'information, l'ouverture des données se place dans la lignée des pratiques anciennes de partage de données en science. Pour Bruno J. Strasser {Strasser:2012kn}, l'émergence des sciences guidées par les données (*data-driven sciences*) trouve ses sources dans les méthodes comparatives des sciences du vivant. Les cabinets de curiosité où étaient entreposées, classifiées et exposées des espèces hétéroclites sont en quelque sorte les ancêtres des bases de données modernes. Cette tradition de collection a abouti au XIXe siècle à l'essor des muséums de sciences naturelles et leurs immenses collections d'espèces. Au début du XXe siècle, cette longue tradition a décliné sous l'assaut de la science

². L'expression anglosaxonne « *knowledge is power* » est souvent employée pour justifier la circulation de l'information. Savoir serait synonyme de pouvoir. Quand j'entends cet argument, je montre souvent un extrait du premier épisode de la saison 2 de la série *Game of Thrones* dans lequel Littlefinger (intrigant conseiller du roi) menace la reine Cersei de dévoiler sa relation incestueuse avec son frère en disant « *knowledge is power.* » Elle demande alors immédiatement à sa garde de l'arrêter et de lui couper la gorge, les interrompant au dernier moment dans leur geste. Elle lui répond alors « *power is power.* » Réduire le pouvoir au savoir, c'est oublier une des caractéristiques essentielles de l'Etat : le monopole de la violence physique légitime cher à Max Weber. J'aime montrer cet extrait pour lancer un débat sur le potentiel émancipateur de l'ouverture des données.

expérimentaliste qui a consacré le laboratoire comme le seul lieu de la recherche scientifique légitime. Le projet de codifier et de numériser l'ADN dans les années 60 a marqué le retour à la tradition comparative de collection et de comparaison des espèces. Strasser a montré qu'une des raisons du succès de Genbank, la première base massive de données en génétique, a reposé sur le partage des données et l'apparence d'ouverture du système. Dans d'autres *data-driven sciences* comme l'astronomie ou la botanique ou la génétique, le partage et la réutilisation de grandes bases de données sont devenus des pratiques courantes. Avec la disponibilité de données de plus en plus importantes, de nouveaux métiers ont émergé comme les *computational scientists* qui analysent les données produites par d'autres. Certains ont vu d'un mauvais oeil ces chercheurs qui publient en leur nom avec les données produites par d'autres, les qualifiant parfois de « chercheurs parasites. »

Devenu la norme dans de nombreuses disciplines, le partage des données doit permettre la reproductibilité des résultats et l'émergence de nouveaux savoirs à partir des données d'un projet de recherche. Une base de données est ainsi devenue une publication scientifique à part entière {Bowker:2000ws}. Dans ce contexte, il est courant d'entendre que les sciences comme les organisations feraient face à un « déluge de données. » Or, ces perceptions apparaissent régulièrement depuis la Renaissance en même temps que les technologies de gestion de la surabondance d'information {Rosenberg:2003wt}. Les articles relatant le « déluge de données » ont aussi charrié l'idée que le caractère massif et exhaustif des bases de données pourrait bouleverser la science. Dans un nouveau paradigme {Hey:2009vh}, les données toujours plus massives n'auraient qu'à « parler » pour interpréter les phénomènes et générer du savoir. Plutôt que de partir d'une hypothèse et de la tester, il suffirait d'explorer les données pour identifier des corrélations, des relations statistiques entre des phénomènes, et comprendre la causalité. Pour Chris Anderson, éditeur de *Wired*, le déluge des données rend la méthode scientifique obsolète. Dans un article célèbre de 2008, il annonçait pas moins que la fin de la théorie :

Le déluge des données rend la méthode scientifique obsolète, l'analyse des motifs et des relations contenues dans les données massives produit intrinsèquement un savoir significatif et éclairé sur des phénomènes complexes. Il y a maintenant une meilleure manière de faire. Les petabytes nous permettent de dire que « la corrélation suffit ». Nous pouvons analyser les données sans hypothèses sur ce qu'elles peuvent montrer.

C'est au même moment que la demande d'ouverture des données est entrée dans

le débat. La possibilité de réutiliser l'ensemble des données du secteur public permettrait de générer des recherches inédites et de créer de nouvelles formes de connaissances.

Les politiques d'open data s'inscrivent donc aussi dans la lignée de pratiques anciennes de partage de données en science. Le terme même « open data » est apparu dans les sciences. Selon Yu et Robinson {Yu:2012wv}, il est employé pour la première fois dans les années 1970 dans des accords que la NASA a signé avec des pays partenaires en vue du partage de données satellitaires. En 1995, la National Academy of Science aux États-Unis a publié un rapport intitulé "*On the Full and Open Exchange of Scientific Data*" qui serait le premier usage public du terme. Notons enfin que c'est dans les sciences que vient la notion de données « brutes », centrale dans les principes de l'open data, désignant les mesures telles qu'elles sont produites par les instruments sans avoir fait l'objet d'un traitement pour les corriger ou les « nettoyer. » Les *Science and Technology Studies* nous ont appris à prendre des précautions avec cette notion souvent associée à un idéal d'objectivité. Bowker {Bowker:2005ud} et Gitelman {Gitelman:2013wl} ont dénoncé un « oxymore et une mauvaise idée » qui répand l'idée que la production de l'information serait transparente et naturelle. Car de fait, l'étude de la production de données scientifiques rend compte de l'impossibilité d'identifier un moment où la donnée aurait connu un état brut, sans traitement {Ribes:2013wf}. Les « données » sont toujours des « obtenues » {Latour:1993ti}. Elles sont le fruit de choix, de théories et d'interventions et ne sont jamais pré-interprétatives comme le laisse penser l'expression « données brutes. »

Les industries de l'information

L'émergence des politiques d'open data doit aussi au développement des industries de la donnée, des entreprises pour lesquels la réutilisation des données constitue le coeur de l'activité. Cette industrie est particulièrement développée aux États-Unis où toutes les données publiques sont dans le domaine public, sans redevances ni restrictions depuis 1988 {Ronai:1997va}. Cette situation est le fruit d'une intense campagne de lobbying par les industriels de l'information qui ont obtenu du gouvernement fédéral que les données soient systématiquement partagées et que les services à valeur ajoutée soient fournis principalement par le secteur privé plutôt que par les acteurs publics. En France, la situation a longtemps été différente avec le succès des services télématiques sur Minitel qui ont dégagé d'importantes sources de revenus dans les années 1980 pour les acteurs publics. Maurice Ronai parle d'une prolifération et d'un « activisme télématique » dans les administrations :

Les administrations exploitent leurs gisements naturels : le vaste domaine des droits et démarches. Les serveurs prolifèrent parfois à l'intérieur d'un même ministère. Des préoccupations nouvelles s'affirment : faire du trafic, générer des recettes. La diversité des pratiques tarifaires, les bricolages contractuels débouchent sur des situations ambiguës : monopoles de commercialisation, accords d'exclusivité, relations de troc, recours à des associations...

Même si le Minitel a progressivement disparu des foyers dans les années 1990 avec le développement d'Internet, ces recettes ont créé un précédent qui a entravé la mise à disposition gratuites en ligne des données publiques. La mission du Premier ministre sur les redevances en matière de données publiques présidée par Mohammed Adnène Trojette a montré que les recettes en France s'élevaient à 34,7 millions d'euros en 2012, en diminution par rapport à 2010 où elles s'élevaient à 52,1 millions d'euros {Trojette:2013wr}.

Les politiques d'open data ont fortement incité à l'abandon de ces redevances. Dans les années 2000, la Commission européenne s'est inspirée du cadre réglementaire états-unien pour promouvoir l'abandon des redevances afin de faire émerger une industrie de l'information de taille comparable à celle des Etats-Unis. En 2003, l'Union Européenne a adopté la directive sur les informations du secteur public (Public Sector Information) pour inciter les États membres à diffuser gratuitement leurs données. La Commission européenne a multiplié les études sur le potentiel économique de la libération des données, évaluant jusqu'à 200 milliards d'euros par an la valeur de leur circulation optimale dans les pays de l'Union {Vickery:2011tk}. Ces promesses de croissance doivent justifier les efforts financiers et humains (suppression des redevances, développement de portails, travail sur les données, communication et événements, recherche et développement) consentis par les États pour ouvrir leurs données. Le pari de l'open data repose sur un retour sur investissement dans la taxation des revenus de la réutilisation dont il est prévu que le montant dépasserait largement le coût de l'ouverture et l'abandon d'éventuelles redevances.

Ces prévisions de croissance se sont transformées en revendication à la suite d'une campagne du quotidien britannique le *Guardian* en 2006 intitulée « Rendez nous les joyaux de la couronne » qui exigeait la gratuité des données « payées par les impôts. » Parmi ces « joyaux », les données de l'institut géographique britannique Ordnance Survey ont été perçues comme les plus prometteuses en terme de

croissance. La tribune insistait sur le fait qu'aucun des services de cartographie numérique comme Google Maps n'était britannique du fait des fortes redevances pour utiliser les données de l'Ordnance Survey. Le Guardian évoquait alors la différence avec la conception américaine :

Aux États-Unis il est considéré que les données collectées avec l'argent du contribuable devraient lui être restituées gratuitement. Et une étude détaillée montre que l'attitude fermée du Royaume-Uni envers ses données signifie que nous perdons des opportunités commerciales et même que nous freinons la recherche scientifique dans des domaines comme le changement climatique³. »

Cette campagne, menée par le principal quotidien du pays, a créé des réseaux et mobilisé des acteurs importants dans le développement de l'open data en Grande-Bretagne.

Rappelons enfin que l'open data est apparu dans le débat public aux alentours de 2007, en même temps que le développement de l'économie des apps avec le lancement des boutiques d'applications sur smartphone (Apple AppStore, Google PlayStore...). Il était espéré une vague de création de services et d'emploi dans le développement informatique avec l'accès ubiquitaire aux services mobiles rendu possible par le smartphone. Dans ce contexte, les données ouvertes étaient souvent envisagées comme un carburant de cette nouvelle économie. Chaque jeu de données ouvertes pouvait potentiellement donner lieu à la création d'une jeune pousse. L'argument de la création d'emplois a joué un rôle important dans le contexte de la crise économique de 2008.

La modernisation des administrations

Au-delà des promesses de croissance, les politiques d'open data reposent aussi sur leur capacité à provoquer des changements dans les rouages de l'administration et ainsi engendrer sa « modernisation ». Les projets d'ouverture des données publiques sont souvent soutenus par les services en charge de la transformation du service public. Ces derniers prévoient que la diffusion des données permettra aux agents de bénéficier des informations des autres services. La réutilisation des données produites dans d'autres divisions de l'organisation permettrait la fin du

³. Charles & Cross (2006). "Give Us Back Our Crown Jewels", *The Guardian*. Accessible en ligne : <http://www.theguardian.com/technology/2006/mar/09/education.epublic>.

« travail en silos », une expression désignant le cloisonnement des services et l'insuffisance des pratiques de partage d'informations. L'ouverture des données publiques offrirait aussi la possibilité de créer des services innovants en utilisant des données qui n'étaient jusqu'alors pas connues. Les administrations en charge de l'efficiencia du service public anticipent aussi que les producteurs vont devenir redevables auprès de leurs collègues de la qualité des données.

Cette modernisation concerne aussi le public lui-même. L'ouverture des données encouragerait la participation des citoyens dans l'élaboration des politiques par la réduction des asymétries d'information. En disposant des données dans leur plus grande précision, les politiques d'open data faciliteraient l'apparition d'une citoyenneté informée au fondement de politiques publiques d'inspiration libérale {Barry:2001vo}. La disponibilité des données publiques permettrait l'émergence de contre-pouvoirs à même de remettre en cause de manière factuelle les politiques publiques en cours d'élaboration ou de mise en oeuvre. Des actions collectives se mobilisent régulièrement pour que les données, les catégories et les indicateurs prennent en compte des situations jusqu'alors occultées ou ignorées {Bruno:2014tm}. Plus largement, l'ouverture de l'information stratégique au public est perçue comme une source d'efficacité et d'innovation des organisations. Les politiques d'open data s'inscrivent dans la continuité des travaux des théoriciens des organisations qui encouragent à diffuser l'information. En particulier, les travaux d'Henry Chesbrough {Chesbrough:2006uz}, à l'origine de la notion d'innovation ouverte, ont répandu l'idée qu'ouvrir l'information stratégique peut constituer une source d'innovation bien supérieure au régime classique du confinement des ressources en interne.

Ces bénéfices internes et externes positionnent les politiques d'open data comme un objet d'intérêt pour la transformation du service public. La Banque Mondiale encourage ainsi financièrement les pays en développement à ouvrir leurs données à la fois pour faire émerger une société civile informée, mais aussi pour « moderniser » le fonctionnement des administrations. L'open data est devenu en quelque sorte un « produit d'exportation » pour les pays occidentaux qui proposent des méthodologies « clé en main » pour transformer l'administration par l'ouverture des données {Birchall:2015km}. Cette « modernisation » s'inscrit aussi dans le contexte de réduction des dépenses de l'État où l'ouverture des données financières permettrait un contrôle de l'action publique qui dépasse les instances dédiées et l'étend à la société civile dans son ensemble. Des recherches

critiques ont associé l'ouverture des données publiques à la privatisation du service public {Bates:2012ue}. Cette critique doit être prise au sérieux : l'ouverture des données peut aussi servir à un désinvestissement de l'Etat dans les services à destination des citoyens. L'État peut se contenter d'ouvrir les données et déléguer au secteur privé la création de services, abandonnant alors le principe d'égalité d'accès au service public. L'éditeur Tim O'Reilly qui a participé à l'organisation de la conférence fondatrice des principes de l'open data (abordée en détail dans le chapitre suivant) a écrit en 2010 l'article très commenté « *Government as a Platform* » {OReilly:2010tc}. Dans une rhétorique libertarienne tirée d'exemples où la société se substitue à l'État, l'article défend une conception de l'État comme une plateforme mettant à disposition de la société civile et des acteurs privés des ressources. Ces derniers ont la liberté de développer des biens et des services finaux à l'aide de ces ressources, une approche qui peut créer les conditions d'un désengagement de l'État.

Tout au long de ce chapitre, nous avons pu voir que l'ouverture des données s'inscrit dans la lignée d'une diversité de filiations, chacune mobilisant une diversité d'arguments et d'acteurs. Cette pluralité des facettes de l'ouverture des données se retrouve tout au long de l'histoire des politiques d'open data que nous allons aborder dans les chapitres suivants, d'abord dans l'émergence des principes de l'open data et ensuite dans sa mise en oeuvre en France. Du fait de ces multiples facettes, l'ouverture des données a pu créer une forme de consensus trans-partisan comme l'a constaté en 2012 Valérie Peugeot en listant les affiliations politiques des élus ayant adopté une démarche d'ouverture des données publiques {Peugeot:2012um}. Selon elle, l'ouverture des données présente les caractéristiques d'une innovation démocratique remarquable du fait de la diversité d'échelles territoriales engagées dans la démarche (villes, départements, régions, États, organisations internationales) et de la rapidité de son adoption par les autorités politiques. Surtout, cette innovation démocratique se distingue selon Valérie Peugeot par son caractère consensuel car « la couleur politique du gouvernement en place au moment de son adoption n'apparaît pas comme un caractère discriminant » et « lorsqu'un changement politique est intervenu, aucun gouvernement n'a cherché à remettre en question ce choix politique. » Cette « convergence en trompe-l'œil » masque selon elle une double lecture de l'ouverture des données :

- une approche libertarienne selon laquelle ouvrir les données constitue une opportunité pour « rendre au marché des fonctions tenues par l'État et donc de contribuer à une réduction de la fiscalité, à une diminution globale du rôle de la puissance publique dans la société »

- une approche issue des biens communs informationnels qui consiste « non pas à défaire l'État mais à en compléter l'action par des formes de gouvernances auto-organisées. »

Pour analyser une politique d'ouverture des données, il faudra donc regarder au-delà de la simple décision d'ouvrir les données et s'intéresser au détail des choix réalisés comme par exemple les types de jeux de données mis à disposition, leur contenu ou les incitations à la réutilisation des données. La question du choix des licences qui encadrent la réutilisation des données constitue le premier de ces choix qui peuvent « teinter » politiquement une politique d'open data en faveur d'une lecture ou d'une autre. C'est cette question de l'encadrement juridique qui a d'abord été traitée dans la définition des grands principes de l'open data.