



N° 4 | 2018

Méthodes et stratégies de gestion de l'information par les organisations : des big data aux thick data

L'horizon d'une culture de la donnée ouverte : de l'utopie aux pratiques de gouvernance des données

Anne Lehmans

Professeur des universités

INSPE

IMS UMR5218 CNRS

University of Bordeaux

Édition électronique :

URL :

<https://revue-cossi.numerev.com/articles/revue-4/1828-l-horizon-d-une-culture-de-la-donnee-ouverte-de-l-utopie-aux-pratiques-de-gouvernance-des-donnees>

DOI : 10.34745/numerev_1600

ISSN : 2495-5906

Date de publication : 04/09/2018

Cette publication est **sous licence CC-BY-NC-ND** (Creative Commons 2.0 - Attribution - Pas d'Utilisation Commerciale - Pas de Modification).

Pour **citer cette publication** : Lehmans, A. (2018). L'horizon d'une culture de la donnée ouverte : de l'utopie aux pratiques de gouvernance des données. *Revue COSSI*, (4).

https://doi.org/https://doi.org/10.34745/numerev_1600

Le développement des open data en France conduit les acteurs à s'interroger sur les stratégies et les pratiques de gestion des données à mettre en place dans les organisations concernées. L'affichage d'une politique d'ouverture des données, dans une logique affirmée de transparence, de participation et d'innovation, est susceptible de bouleverser les routines dans les modes de gestion et de contrôle de la circulation de l'information. Les principes et les formes de gouvernance des données font l'objet d'une réflexion renouvelée, l'ouverture des données faisant office de catalyseur pour introduire un principe de décision partagée dans le cycle de vie de la donnée. Un projet de recherche sur la culture des données, partant d'une enquête qualitative sur les pratiques de gestion des données, montre que, face aux demandes, aux risques et aux avantages perçus dans l'agenda de l'ouverture et de la diffusion des données ouvertes, des stratégies variées de gouvernance des données s'installent, avec des effets sur le management de l'information et la gestion des connaissances.

Mots-clefs :

Open data, Données ouvertes, Gestion des données, Gouvernance des données, Cycle de vie de la donnée

Abstract : The development of open data in France leads stakeholders to question the strategies and data management practices to be implemented in organizations. The display of an open data policy, in a logic of transparency, participation and innovation, is likely to upset the routines in the management and control of the flow of information. The logics and forms of data governance are subject of a renewed reflection. Open data is a catalyst to introduce a shared decision-making principle into the data lifecycle. A research project on data literacy, based on discourses analysis and on a qualitative survey on data management practices, shows that, faced with the demands, risks and benefits perceived in the agenda of open data, various strategies are taking hold, with effects on the management of information and knowledge.

Keywords : open data, data governance, data management, data life cycle

INTRODUCTION

Nous assistons, depuis plusieurs années, au phénomène d'explosion des "data" ou "données", décliné sous de multiples formes, big data (données massives), open data (données ouvertes), linked data (données liées). Si la question des données a d'abord

concerné la sphère professionnelle, elle a rapidement opéré une transformation des dynamiques sociales et culturelles, faisant émerger de nouvelles problématiques économiques, sociales et politiques, et la nécessité d'une culture numérique. En effet, les données ont à voir avec la construction et la diffusion de la connaissance. L'économie de la connaissance repose sur un modèle de valorisation de l'information à partir des données disponibles pour la production de savoirs d'innovation. Cependant, ce modèle, basé sur l'appropriation privée des idées et de la création, est interrogé depuis longtemps par la société civile, les entreprises, et par des chercheurs dans plusieurs disciplines, spécialistes de l'information, économistes, informaticiens, géographes, politologues, designers. De nouvelles dynamiques territoriales émergent autour des thèmes de l'open-innovation, de la participation, de l'économie sociale et solidaire, du partage de biens matériels (Demailly, Novel, 2014) et surtout de la politique d'ouverture des données (*open data*) officialisée en 2011 en France avec la mission Etalab, à laquelle une grande partie de la loi pour une république numérique d'octobre 2016 est consacrée.

Les données ouvertes (*open data*) désignent les données collectées par des organismes publics ou privés chargés d'un service public, ou les citoyens, et mises à disposition en format numérique sur des plateformes nationales ou locales permettant leur accès et leur réutilisation. Essentiellement quantitatives, elles peuvent être traitées, triées, croisées, exploitées, et visualisées à des fins de développements, de recherche, ou d'enseignement. Elles ont des caractéristiques spécifiques dans l'économie de l'information, puisqu'elles sont variées, souvent gratuites, ouvertes et exploitables, mais nécessitent des traitements informationnels pour être intégrées aux activités des professionnels ou du public. Si la question des données ouvertes évolue rapidement dans les agendas politiques, du local à l'international, la réalité des pratiques en France, si on les compare aux pratiques canadiennes, par exemple, reste hétérogène. L'affichage d'une politique d'ouverture des données est susceptible de bouleverser des formes de routine dans la gestion et le contrôle de la circulation de l'information dans les organisations. Ce sont, en particulier, les modes de gouvernance des données qui font l'objet d'une réflexion renouvelée, l'ouverture des données faisant office de catalyseur pour introduire un principe de décision partagée. On voit bien surgir la question de l'épaisseur de la donnée, qui n'existe pas en dehors de processus de gestion, de médiation et de médiatisation. Elle nécessite de nouvelles formes de gouvernance qui font entrer les usagers dans l'écosystème, entre l'accès et l'utilisation.

Une recherche sur les politiques et les pratiques de médiation et d'éducation autour des données ouvertes dans les collectivités territoriales, visant à comprendre les circuits de décision et les représentations^[1], révèle que l'ouverture des données et les demandes sociales liées à la mise à disposition de jeux de données déclenchent des questions nouvelles. Les services chargés de la gestion des données ouvertes répondent de façon pragmatique aux demandes en s'organisant en fonction des compétences disponibles et des soutiens qu'ils trouvent au sein de la collectivité, du côté de la hiérarchie ou du côté des politiques. C'est ce que montre une enquête qualitative qui a permis de repérer des chaînes d'usages, d'analyser les discours des acteurs dans les collectivités et les valeurs qu'ils portent, pour caractériser des dynamiques de gouvernance. A partir de la

problématique de la recherche, présentée dans un premier temps, on verra que ces dynamiques répondent à des injonctions contradictoires et relèvent de pratiques de bricolages, de détournements des routines qui contribuent à modifier les processus de production et de gestion des données, à long terme à les enrichir.

PROBLÉMATIQUE : LE LIEN ENTRE GOUVERNANCE ET CULTURE DES DONNÉES

L'ouverture des données publiques confirme le mouvement démarré en France en 1978 par la loi créant la commission d'accès aux documents administratifs. Ce mouvement s'est accéléré depuis la création d'un grand portail data.gouv.fr en 2011 et de la mission Etalab, les lois Valter (sur la gratuité des informations publiques) en 2015 et Lemaire (loi pour une République numérique) en 2016. Dann Goncalves et Samuel Rufat (2016) rappellent que l'expression « open data » vient de la recherche publique environnementale et que le concept a été créé pour répondre à un triple besoin : partager les données pour favoriser une réflexion globale, baisser le coût de ce partage, et favoriser la participation. Du point de vue des politiques publiques, *l'open data* repose sur les principes de transparence, de participation citoyenne et de modernisation de l'action publique par la collaboration entre les institutions et les citoyens (El Hachani, 2015 : 5) et le souci de "réutilisation intelligente des données indépendamment de leur contexte numérique d'origine" (Noyer, Carmes, 2012 : 2). Le rapport du Secrétariat général pour la modernisation de l'action publique de 2015 sur la gouvernance de la donnée ajoute que "la mise à disposition de données libres et ouvertes, et leur réutilisation, génèrent de la valeur économique et sociale, par le biais de cinq mécanismes : la réduction des coûts de transaction, l'innovation, la réduction des asymétries d'information, la collaboration et les boucles de rétroaction" (p. 21).

Si la transparence et la participation sont au coeur du discours politique sur l'ouverture des données, leur mise à disposition et leur accessibilité publique s'inscrivent dans une logique de "communs de la connaissance" (Hess et Ostrom, 2007) qui a introduit une possibilité de "gouvernance" ouverte et un nouvel enjeu de société. Cependant, pour que les données ouvertes puissent être utilisées, leur traitement nécessite des compétences informationnelles, documentaires mais aussi techniques (informatiques) et soulève des questions socio-économiques et politiques. Il s'agit non seulement d'assurer l'accessibilité de méthodes et savoir-faire pour créer et exploiter les données ouvertes, mais aussi de construire activement une démocratie participative dans le contexte émergent des "villes intelligentes" tout en favorisant le développement économique régional, notamment dans le cadre de l'économie sociale et solidaire.

Le projet de recherche sur la médiation des données pour l'éducation met en relation la question de la formation aux données avec celle de leur gouvernance dans les organismes publics et privés, qui appelle de nouvelles compétences et connaissances (Capelle, Lehmans, Liquète, 2017). Il concerne la façon dont se construit l'offre de jeux de données au sein de l'écosystème informationnel régional, et les liens à nouer avec les professionnels de l'éducation, de la formation et les associations pour former aussi

bien les jeunes que tous les acteurs de la société aux données et à leur appropriation. En préalable, un point important est celui de l'efficacité des politiques d'ouverture des données, à travers l'instrumentation de l'action publique et la dynamique impulsée par le répertoire d'instruments utilisés dans la mise en oeuvre de la politique d'ouverture des données. L'innovation, liée à la modernisation de l'action publique, peut être considérée comme une forme d'instrumentalisation de la gouvernance (Lascombes, Simard, 2011) à travers un récit performatif qui a des effets réels sur le cycle de vie de la donnée dans les organisations. Mais l'accélération des politiques publiques ne va pas nécessairement de pair avec leur mise en oeuvre dans les organisations, parce que l'ouverture des données représente un changement profond, du point de vue du management de l'information notamment.

Dans le domaine des sciences de l'information et de la communication, les chercheurs travaillent depuis plusieurs années sur les problématiques d'ouverture des données publiques des collectivités territoriales et de l'Etat. La diffusion des données publiques est souvent présentée comme faisant partie intégrante de la mission de service public (Labelle, Le Corff, 2012). Dans le même temps, les portails *open data* des collectivités territoriales présentent une « uniformisation des modes de communication des informations » (Idem). La mise à disposition des données devrait permettre une réutilisation et une réinterprétation dans un cadre personnalisé. Mais on voit émerger des questions auxquelles la recherche n'a encore que très partiellement répondu : le lien entre les cultures de travail, les identités professionnelles et les usages numériques des réseaux, les nouveaux collectifs de travail et les communautés épistémiques (Crowston, Jullien, Ortega, 2013), la dimension participative voire militante (Cardon, Granjon, 2010) des usages numériques.

Dans toutes ces questions, c'est la gouvernance de l'information qui est en jeu. Le terme de gouvernance a longtemps connu plus de succès dans les pays anglo-saxons et au Canada qu'en France, où il a été relégué pour son caractère polysémique. Il fait néanmoins une apparition remarquable depuis quelques années dans les documents institutionnels produits par les administrations, dans lesquels il fait l'objet d'analyses et de propositions. La gouvernance, par rapport au gouvernement, a une connotation moins institutionnelle et plus dynamique : quand on parle de gouvernance, on s'intéresse aux processus réels de prise de décision plus qu'aux structures formelles, ainsi qu'à la diversité des acteurs et des modes de participation possibles (Paye, 2005). Le Secrétariat général pour la modernisation de l'action publique a remis en 2015 un rapport sur « La gouvernance de la donnée » qu'il définit comme "l'ensemble de principes et de pratiques qui visent à assurer la meilleure exploitation du potentiel des données (p. 48)." Le terme est ici très générique, se référant à des principes, donc un cadre, mais aussi des pratiques. Le rapport pointe les limitations dans le potentiel des usages des données, liées à la méconnaissance, à l'imperméabilité de la culture administrative par rapport aux dynamiques de coopération et de participation. Il préconise de partir de développements concrets, de cartographier les données disponibles, de faire évoluer les systèmes d'information et de décloisonner les administrations. Les questions qui surgissent concernent les modes de fonctionnement politique et bureaucratique des organisations, leur relation avec le public et les façons

de l'intégrer à la prise de décision, les compétences professionnelles, et, de façon générale, l'écosystème informationnel. La gouvernance « associe négociations, prises de décision et émergence de consensus» (Schafer et Le Crosnier, 2011) à travers des formes de participation et d'engagement qui se structurent par le travail sur les données. Le concept de gouvernance ouverte lie les principes de transparence et de participation dans une relation complexe intégrant visibilité de l'action publique et développement des moyens d'expression, voire d'action, des citoyens. Son analyse doit tenir compte des "comportements organisationnels" qui peuvent constituer des entraves à l'ouverture (Pasquier, Villeneuve, 2007). Celle-ci a pour corollaire un principe de confiance visant à maintenir un équilibre et éviter que certains s'approprient des éléments d'information au détriment d'autres (Maurel, Chebbi, 2012 : 85). Mais les enjeux politiques de l'ouverture des données sont eux-mêmes fortement corrélés à des questions économiques dans le cadre global du big data.

Les objectifs de cette première recherche sont d'identifier les représentations des usages potentiels et réels des données ouvertes. Le projet de recherche a permis d'enquêter sur le terrain de collectivités locales productrices de données auprès d'une quinzaine d'acteurs en Aquitaine, des médiateurs et des utilisateurs finaux des jeux de données ouvertes. L'équipe de recherche a mené des observations participatives avec des séances de travail sous forme d'ateliers et de focus groupes afin de comprendre les contraintes et les freins éventuels au développement d'une culture des données, ainsi que les conceptions de cette culture qui circulent dans les organisations. Plusieurs critères ont été analysés : l'offre de données, les stratégies de mise à disposition, les moyens de la gouvernance sur le plan des ressources humaines, techniques et organisationnelles, les représentations des usages sociaux possibles, les ressources et les stratégies de médiation et de formation dans l'écosystème informationnel mis en place par les acteurs, les stratégies d'évaluation qu'ils ont choisies. En parallèle des observations, l'équipe a mené des entretiens compréhensifs afin d'identifier et de caractériser le discours des acteurs de terrain pour comprendre les objectifs qu'ils se fixent, leur représentation de la chaîne de production et de diffusion des données dans la sphère publique actuelle. Il s'agissait de comparer les représentations d'acteurs entre des secteurs d'activités et des échelles géographiques différents, et de saisir l'évolution de leurs pratiques déclarées sur leurs terrains. L'enquête révèle un ensemble de bricolages, le déploiement de tactiques, au sens des « arts de faire » de Michel de Certeau (1990), dans l'épaisseur concrète des pratiques informationnelles et professionnelles. Ces bricolages ne mettent pas en question les stratégies qui soutiennent les politiques d'ouverture des données, mais en révèlent les contradictions possibles.

BRICOLAGES IDEOLOGIQUES : LES CONTRADICTIONS DES SYSTEMES DES VALEURS, ENTRE INJONCTION D'INNOVATION ET TRANSPARENCE

Les politiques d'ouverture des données mises en place ne rencontrent pas encore

l'adhésion des organisations pour un passage à l'échelle globale, malgré l'existence de nombreuses expériences locales très positives. Le discours de la transparence relève en grande partie d'une utopie (Gallot, Verlaet, 2016) et d'une mise en scène de l'information. La transparence de l'action publique peut se décréter, mais l'ouverture et la mise à disposition de données nécessitent un réel effort, une volonté politique et sa mise en oeuvre par des services, donc des moyens. La mise à l'agenda politique et budgétaire de l'ouverture des données reste très inégale et hétérogène, soit parce que les enjeux de l'ouverture des données ne sont pas encore compris des collectivités territoriales et notamment des élus, car trop techniques et complexes, soit parce que l'ouverture met en danger l'image de la collectivité ou de l'administration et se heurte à de très fortes résistances (Goëta, Mabi, 2014). Par ailleurs, transparence et participation supposent que les citoyens sont capables de s'emparer des données, qu'ils ont la culture nécessaire pour en faire des usages variés. Les entreprises peuvent disposer de ces compétences qui nécessitent des moyens financiers pour développer, par exemple, des programmes (API), mais pour les citoyens, les usages potentiels sont beaucoup plus complexes. Ainsi, dans la double cible visée par l'ouverture des données, les entreprises et les citoyens, l'une est très dominante. Les objectifs du projet *d'open government* porté par des activistes du libre comme Lawrence Lessig, Tim O'Reilly ou Aaron Swartz, dans la réunion de Sébastopol, (Californie) en 2007, visant le développement d'un pouvoir de contrôle citoyen sur le gouvernement et la correction des asymétries d'information paraissent, dix ans plus tard, très lointains encore.

En revanche, les objectifs économiques de soutien à l'innovation par l'ouverture d'une manne informationnelle en direction des entreprises sont bien plus clairs et faciles à atteindre, et l'on voit se développer de nouveaux services à partir des données récupérées. L'innovation concerne essentiellement le monde économique. Finalement, le risque est important de voir les politiques d'ouverture renforcer le pouvoir de ceux qui ont les moyens économiques d'en faire usage en laissant de côté les citoyens. C'est le point de vue de certains et le risque identifié par les collectivités territoriales elles-mêmes (Gurstein, 2011). Les données ouvertes illustrent très clairement la nécessité de distinguer l'accès et l'usage, la garantie de l'accès n'étant absolument pas une condition suffisante pour que des usages existent. Michael Gurstein propose un modèle d'usage des données qui prend en compte l'accès à internet, à des machines, les compétences techniques en informatique, le contenu et le format des données, les connaissances pour les interpréter, l'accompagnement et la gouvernance, qu'il définit comme les moyens financiers, juridiques, réglementaires et politiques nécessaires à la mise en oeuvre d'une ouverture qui vise les usages réels et pas seulement les déclarations ou les usages de grands groupes économiques. La gouvernance, sous toutes ses dimensions organisationnelle, normative, technique, politique, renvoie à des valeurs de référence qui sont à la base difficilement compatibles.

En outre, une question importante et récurrente est celle du risque pour les citoyens. En effet, dans les croisements de grandes séries de données, avec l'expertise des analystes et la mise en place d'algorithmes, donc la mise en oeuvre de moyens que seuls les Etats ou les grandes entreprises du GAFAM peuvent actionner, il est tout à fait possible de produire de l'information stratégique (avec des risques sanitaires par

exemple), bien-sûr, mais surtout personnelle, même si les données ouvertes sont anonymes. Ces risques n'échappent pas aux acteurs, qui observent la disparition des derniers tabous de réutilisation de clés d'identification, sous couvert d'amélioration et d'optimisation du service public. Le risque pour les libertés individuelles n'est donc pas négligeable.

Sur le plan des principes, une antinomie se révèle entre les discours de la participation et ceux de l'innovation et de l'efficacité. Les valeurs de participation appellent des modes de gouvernance complexes qui associent les citoyens dans le processus de diffusion et d'usage des données, au prix d'efforts de standardisation, de description, de communication, de formation et d'accompagnement très importantes, dans une temporalité lente, par la constitution d'un réseau d'acteurs. Les valeurs d'innovation appellent une action beaucoup plus rapide, ciblée sur les besoins des entreprises qui prennent en charge les compétences nécessaires. A cette première contradiction dans les objectifs mêmes de l'ouverture des données s'ajoutent les difficultés organisationnelles.

BRICOLAGES ORGANISATIONNELS ET DISCOURS DE L'EFFICACIE : REVOIR LA GESTION DES MÉTIERS AU DÉTOUR DES DONNÉES

Face à l'injonction politique et juridique, aux demandes provenant de la société civile ou aux choix délibérés d'ouverture des données, les collectivités ont dû mettre en place des procédures de gestion dans l'urgence. Sur le plan des discours, elles sont contraintes à l'affichage d'une dynamique à laquelle elles ne sont pas toujours préparées et qui nécessite des formes d'acculturation et des réorganisations des procédures et des modes de travail. Les problématiques d'accessibilité et d'usabilité viennent perturber les fonctionnements de services cloisonnés dans la gestion des données (autour des directions des systèmes d'information), de l'information et des documents, et mettent en relief, sinon des dysfonctionnements, du moins des résistances au changement qui les placent en porte-à-faux avec les services chargés de la communication, d'une part, les politiques, d'autre part. Du point de vue de la gouvernance, l'impulsion et le dynamisme des politiques d'ouverture des données nécessitent des investissements importants. Les contenus et les modalités de cette gouvernance dépendent souvent de l'histoire de l'organisation. Les acteurs témoignent des décalages des temporalités. La décision de mettre en place un portail a été prise, par exemple, à la suite de la publication d'un article de presse pour une collectivité, qui déclenche une décision politique à laquelle les services ne sont pas préparés, dans une logique communicationnelle qui appelle un affichage et un positionnement dans le jeu politique local. Ailleurs, ce sont les directions des systèmes d'information qui donnent l'impulsion, avec une coloration technocentrée et un public cible de développeurs. Le suivi des politiques d'ouverture repose sur les stratégies qui peuvent évoluer en fonction des services qui s'emparent de la question. Dans une collectivité, le caractère tardif et politique de l'ouverture accompagne le recrutement de responsables plutôt non

techno-centrés, et finalement sensibles aux enjeux de médiation, de formation et d'accompagnement du public. Pour certains acteurs, le problème et le retard pris par les collectivités territoriales vient du fait que l'offre précède le besoin, qu'elle est projetée, avec des présupposés parfois contradictoires, alors que, dans le cas d'une plateforme d'information géographique, c'est le besoin des acteurs qui donne l'impulsion et contraint les acteurs, même privés, à céder du terrain informationnel pour améliorer leur propre réservoir.

Quelle que soit l'impulsion de départ, les services sont organisés en silos fonctionnels et en "métiers" communiquant peu entre eux. Ils se sont vus confrontés à un enjeu commun et transversal, qui les contraint à reconsidérer, d'une part l'organisation globale des procédures de collecte et de gestion des données, d'autre part les compétences métiers, et par dessus tout, le partage d'une vision commune de la donnée. Les acteurs témoignent du fait que cette cohérence est indispensable à l'usabilité des données. La diffusion des données environnementales implique, par exemple, que les services chargés de leur collecte connaissent la structuration de la base de données, et que la direction du système d'information soit consciente des contraintes et des spécificités liées à la collecte. En l'absence de cette coopération et de cette connaissance mutuelle, les données mises à disposition du public sont inexistantes, lacunaires, "sales", inutilisables en l'absence d'un travail de nettoyage trop coûteux pour la plupart des usagers. Du point de vue de l'image de l'organisation, elles sont contre-productives.

A l'intérieur même des organisations comme les collectivités territoriales, la culture de la donnée se construit très diversement, en fonction de l'organisation des services, des prises de position des acteurs et de leur capacité à se faire entendre en bousculant les équilibres mis en place notamment autour des systèmes d'information (Lehmans, 2017). Les élus sont souvent peu mobilisés parce que la question des données ouvertes est technique et complexe. L'impulsion de l'ouverture peut venir de services de communication, des services informatiques, des services techniques, des directions générales ou, beaucoup plus rarement, des élus. Les services qui assurent des missions techniques dans le cycle de vie de la données ne partagent pas nécessairement des représentations positives de l'*open data*. Ils y voient des contraintes nouvelles, des risques de bouleversement des procédures de travail en place, de remise en cause de leur mission par la publicité faite aux données statistiques, par les nécessités de standardisation ou de description dans les métadonnées, mais rarement des objectifs valorisants. Quand l'histoire de l'ouverture des données se fait sur ce modèle, la construction d'une culture de la donnée est longue et vient plutôt de l'émergence d'un demande sociale ou politique. Quand l'impulsion vient de l'intérieur, par exemple de la direction des services d'information, le cheminement est différent et la construction d'une culture partagée, parfois sur fond de compétition entre les services, plus rapide.

Les organisations ont souvent mis en place des dispositifs de médiation pour faire évoluer les pratiques, notamment à travers un fonctionnement « en mode projet ». Au-delà des compétences techniques, surtout informatiques, mais aussi liées aux méthodes de collecte et de relevé des données, c'est véritablement un ensemble de pratiques

informationnelles, de représentations techniques et sociales et de savoirs partagés qui est en jeu, ce qui définit le cœur même d'une culture et de la construction de l'identité professionnelle. Parmi ces pratiques, les processus de normalisation sont essentiels.

BRICOLAGES NORMATIFS : GOUVERNANCE ET CULTURE DE L'INTEROPÉRABILITÉ

L'ouverture des données, dans les principes, relève d'un projet ou d'un discours de partage, de transparence, de diffusion démocratique et égalitaire des bases de la connaissance. Or, si l'on considère la réalité des usages des données ouvertes au niveau d'une région par exemple, on constate que les réutilisations réelles et projetées par les diffuseurs concernent essentiellement le marché et les entreprises. Les acteurs des données ouvertes hésitent entre un discours d'ouverture vers les citoyens et de partage démocratique, et une stratégie de développement économique qui vise plutôt les entreprises, notamment celles qui vont créer des applications. Ces hésitations ont des répercussions sur les priorités thématiques d'ouverture, ainsi que sur les normes, les standards et surtout sur les licences.

Concernant les standards, leur respect est la condition de possibilités de partage et d'interopérabilité. La standardisation peut permettre de cristalliser la gouvernance des données. Dans le domaine de l'information géographique, la directive INSPIRE^[2] de la Direction générale de l'environnement de la Commission européenne datant de 2007, *"vise à établir en Europe une infrastructure de données géographiques pour assurer l'interopérabilité entre bases de données et faciliter la diffusion, la disponibilité, l'utilisation et la réutilisation de l'information géographique en Europe"* (Conseil national de l'information géographique). Dans ce cas, c'est au niveau politique et à l'échelon européen que des standards ont été imposés afin d'obtenir une information structurée, mise à jour et partageable, qui a permis le développement d'outils comme le Géocatalogue et le Géoportail et le fonctionnement de plateformes comme PIGMA (plateforme d'information géographique mutualisée en Aquitaine). L'application de cette directive oblige, dans l'organisation chargée de l'information géographique, les thématiciens et les informaticiens à adopter des normes, des langages, des pratiques communs visant l'interopérabilité des données. Pour les acteurs eux-mêmes, il s'agit d'une véritable culture, technique et normative, qui a donné aux services d'information géographique une longueur d'avance sur les collectivités nouvellement arrivées dans l'*open data*, peu sensibilisées aux nécessités fortes de standardisation et à des outils techniquement complexes, qui appellent de nouvelles compétences. Cette standardisation est la condition de la mutualisation et les médiateurs de données la mettent au cœur de leur activité en prenant en charge la formation des services producteurs aux standards et à la structuration des données.

La question des licences reflète les hésitations qui marquent la naissance de l'*open data* à travers trois visions sensiblement différentes : celle de l'Open Knowledge Foundation qui définit l'*open data* à partir des trois critères de la disponibilité et l'accès, la réutilisation et la redistribution, la participation universelle, la seconde qui « vise à une

ouverture toujours plus large des données » (Goetha, 2012), celle des libertaires qui préconise la « *Raw Data now !* » de Tim Berners-Lee. Dans une logique de communs de la donnée, les licences devraient privilégier le partage à l'identique, qui est contraire aux intérêts commerciaux des entreprises et un frein à l'utilisation des données ouvertes. La plupart des collectivités territoriales et *Open Data France* ont donc renoncé à adopter des licences trop strictes (OdBL) imposant un partage à l'identique, et fait le choix pragmatique de permettre aux entreprises de privatiser les services produits à partir des données ouvertes (licence OL créée par Etalab). L'adoption de ces licences vise l'échange né de la réciprocité du besoin d'information. Cette réciprocité s'incarne dans le principe d'interopérabilité.

La problématique de l'*open data* a fait émerger une culture nouvelle qui ne porte plus sur des domaines d'information spécialisés mais sur une base nécessairement commune de traitement de la donnée pour offrir son accès, son usabilité, sa pérennisation, et finalement son épaisseur. Les données de faible qualité sont inutilisables dans une logique de partage, et l'amélioration de la qualité des données apparaît comme une tâche complexe, collective, qui contraint les services à travailler de concert, souvent avec un accompagnement qui vient d'un médiateur ou d'un animateur. Les données de faible qualité sont celles qui sont périmées, non structurées et/ou pas lisibles par machine. La métadonnée apparaît ainsi comme un élément pivot de la culture technique de la donnée, cristallisant des enjeux de compétence informationnelle. L'accompagnement à la création et la structuration de métadonnées est essentiel (Ben Abdallah, 2016). Le schéma DCAT qui vise à normaliser la description des catalogues d'informations publiques, maintenu par le W3C, organisé autour de l'utilisation des concepts de catalogue, de jeux de données, de vocabulaire contrôlé, de catégorie, d'agents et de ressources, peut être mise en avant. Son usage repose sur un travail d'indexation à l'aide d'un vocabulaire contrôlé. Les métadonnées sont ainsi un enjeu politique car elles permettent la contextualisation et l'interprétation des jeux de données, essentielles pour comprendre l'information. Mais cette dimension documentaire n'est pas prise en charge par les services informatiques qui considèrent qu'elle n'entre pas dans leur cœur de métier.

CONCLUSION : ENJEUX ET PERSPECTIVES

Au-delà de la prise en considération des risques et de contraintes techniques et économiques fortes, la gouvernance des données ouvertes interroge la possibilité d'une participation citoyenne à la gestion du cycle de vie de l'information. Mais cette question du cycle de vie de l'information, au cœur des principes de gouvernance de l'information dans les entreprises, n'apparaît pas de façon évidente dans les discours, dans les pratiques et dans les normes concernant les données ouvertes. Elle appelle des formes d'explicitation des décisions et des procédures et de responsabilisation aussi bien des acteurs de l'ouverture que des usagers des données. Elle est complexe, en rupture par rapport à la culture administrative française centralisée, et croise des questions qui relèvent de l'expertise technologique. Des types de légitimité contradictoires sont en jeu, l'une axée sur la démocratie et le dialogue, l'autre sur l'expertise technique et

l'efficacité.

La gouvernance des données ouvertes est intimement mêlée à la construction d'une culture plurielle (Liquète, 2017) de la donnée, à la fois en interne dans les organisations, et du côté des usagers. Elle soulève des questions culturelles liées à des savoirs, des pratiques, des valeurs et des représentations partagés. Si la participation est valorisée, cette dimension culturelle devient centrale. La capacité à se saisir des données ouvertes (de même que des données personnelles) dépend de la formation des individus, et de la culture qu'ils ont pu construire dans leur parcours de formation et des espaces sociaux associatifs, familiaux, amicaux ou professionnels. L'efficacité d'une gouvernance ouverte ne peut faire l'économie de considérations éducatives, pas plus que l'éducation ne peut ignorer la question de l'ouverture des données.

BIBLIOGRAPHIE

Ben Adballah, N. (2016). Schéma de métadonnées pour les Open data : quelles fonctions et quelle harmonisation ? In Paquienséguy F. (dir.), *Open data. Accès, territoires, citoyenneté : des problématiques info-communicationnelles*. Paris. Editions des archives contemporaines.

Bonnaccorsi, J. (2016). Conclusion générale. In Paquienséguy F. (dir.), *Open data. Accès, territoires, citoyenneté : des problématiques info-communicationnelles*. Paris. Editions des archives contemporaines.

Capelle, C., Lehmans, A., Liquete, V. (2017). De la visibilité à la médiation : l'open data et ses usages en éducation. *Colloque international « Big data et visibilité en ligne : un enjeu pluridisciplinaire de l'économie numérique »*, Novembre, Fort de France, Martinique. [halshs-01633284]

Cardon, D. (2014). Apprendre/désapprendre : sur la ligne de crête des apprentissages numériques. *InternetActu.net*, 10/01/14 <http://www.internetactu.net/2014/01/10/apprendredesapprendre-sur-la-ligne-de-crete-des-apprentissages-numeriques/>.

Cardon, D., Granjon, F. (2010). *Médiactivistes*. Paris: Presses de Sciences Po (P.F.N.S.P.).

Certeau, M. de (1990). *L'invention du quotidien. 1. Arts de faire*. Paris, Gallimard.

Chartron, G., Broudoux, E. (2015). Enjeux géopolitiques des données, asymétries déterminantes. Document numérique et société. In *Open Data, big data : quelles valeurs ? Quels enjeux ?*. Bruxelles, De Boeck, 67-83.

Chignard, S. (2012). *Open data, comprendre l'ouverture des données publiques*. Paris, FYP.

Crowston, K., Jullien, N., Ortega, F. (2013). Sustainability of Open Collaborative

Communities: Analyzing Recruitment Efficiency. *Technology Innovation Management Review*, 3(1), 20-26. URL : <http://timreview.ca/article/646>

Demilly, D., Novel, A.-S. (2014). Économie du partage : enjeux et opportunités pour la transition écologique, *IDDRI Studies*, 03/14.

El Hachani, M. (2015). Open data, collectivités et usagers : une dynamique en question. In Paquienéguy F. (dir.), *Open data. Accès, territoires, citoyenneté : des problématiques info-communicationnelles*. Paris. Editions des archives contemporaines. 1-23.

France, Secrétariat général pour la modernisation de l'action publique (2015). La gouvernance de la donnée, URL : http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/rapport_agd_deceembre2015.pdf.

Gallot, S., Verlaet, L. (2016). La transparence : l'utopie du numérique ? *Communication & Organisation*, 49(1), 203-217.

Goëta, S., Mabi, C (2014). L'open data peut-il (encore) servir les citoyens ? », *Mouvements*, 3-79, 81-91. URL : [http://www.cairn.info.docelec.u-bordeaux.fr/revue-mouvements-2014-3-page-81.htm](http://www.cairn.info/docelec.u-bordeaux.fr/revue-mouvements-2014-3-page-81.htm)

Goncalves, D., Rufat, S. (2016). Open data et droit de la donnée : les collectivités à l'épreuve des réglementations européennes. *Cybergeog : European Journal of Geography* [En ligne], Science et Toile, document 787, mis en ligne le 17 août 2016. URL : <http://cybergeog.revues.org/27750>.

Gurstein, M.B. (2011). Open data: empowering theempowered or effective data use for everyone?. *First Monday*.16(2).URL : <http://firstmonday.org/article/view/3316/> 2764.

Hess, C., Ostrom, E. (dir.) (2007). *Understanding Knowledge as a Commons : From Theory to Practice*. Cambridge, The MIT Press.

Labelle, S., Le Corf, J. (2012). Modalités de diffusion et processus documentaires, conditions du « détachement » des informations publiques. Analyse des discours législatifs et des portails open data territoriaux. *Les Enjeux de l'information et de la communication*, 13/2,(2), 209-59.

Lehmans, A. (2017). Données ouvertes et redéfinition de la culture de l'information dans les organisations: Vers une culture de la donnée. *Communication & Organisation*, 51,(1), 15-26.

Liquète, V., de Lourdes, Oliveira, I., Salgueiro Marques, Â. (2017). Les nouvelles cultures de l'information dans les organisations. *Communication & Organisation*, 51,(1), 5-11.

Maurel, D., Chebbi, A. (2012). La perception de la confiance informationnelle. *Communication et organisation*, 42, URL : <http://communicationorganisation.revues.org/3864>.

Noyer, J.M., Carmes, M. (2012). Le mouvement " Open Data " dans la grande transformation des intelligences collectives et face à la question des écritures, du web sémantique et des ontologies. URL : http://archivesic.ccsd.cnrs.fr/sic_00759618.

Pasquier, M., Villeneuve, J-P. (2007). Les entraves à la transparence documentaire. Établissement d'une Typologie et analyse des comportements organisationnels conduisant à empêcher ou à restreindre l'accès à l'information. *Revue Internationale des Sciences Administratives*, 73(1), 163-180

Paye, O. (2005). La gouvernance : d'une notion polysémique à un concept politologique. *Études internationales*, 36-1, mars, 13-4. URL : <http://www.gouvernement.fr/sites/default/files/liseuse/6252/master/projet/Rapport-au-Pr-emier-ministre-sur-la-gouvernance-de-la-donn%C3%A9e-2015.pdf>.

Schafer, V., Le Crosnier, H. (2011). *La Neutralité de l'Internet : une question de communication*. Paris, CNRS éditions.

[1] Le projet IDEX Médiation et médiatisation des données pour l'éducation a été conduit par l'équipe RUDII de l'IMS, UMR5218. Voir <https://dcultures.hypotheses.org/>

[2] Directive 2007/2/CE du Parlement Européen et du Conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE), http://cnig.gouv.fr/?page_id=8991