

Les réinventions de la démocratie à l'aune de l'ouverture des données : du discours de la participation aux contraintes de la gouvernance

Article inédit, mis en ligne le 15 novembre 2018.

Anne Lehmans

Anne Lehmans (anne.lehmans@u-bordeaux.fr) est maître de conférences en sciences de l'information et de la communication à l'Université de Bordeaux (ESPE d'Aquitaine). Elle appartient à l'équipe de recherche RUDII (Représentations, usages, développements et ingénieries de l'information) du groupe Cognitique à l'IMS (Intégration matériaux-système, CNRS-UMR5218).

Plan de l'article

Introduction

Gouvernance ouverte et lisibilité de l'espace public de la donnée

Formatage et utilisabilité de l'espace public de la donnée

Cultures de la donnée, démocratisation de l'espace public de la donnée et communs de la connaissance

Conclusion

Références bibliographiques

RÉSUMÉ

Une recherche sur les politiques et les pratiques de médiation, de valorisation et d'éducation autour des données ouvertes (open data), révèle que celles-ci peinent à trouver leur place dans l'espace public mais sont pourtant porteuses d'une interrogation fondamentale sur le rôle des données dans le fonctionnement de la démocratie contemporaine. L'espace public des données ouvertes est déterminé dans des lieux de négociation et selon des arrangements peu lisibles. Les dispositifs qui permettent l'accès à ces données répondent à une logique de gouvernance administrative et économique plus que démocratique. Ainsi, pour que les données ouvertes participent d'une réinvention de la démocratie dans le contexte du *big data*, les pratiques informationnelles et cognitives qui permettent les usages des données nécessitent des formes d'acculturation qui sont complexes.

Mots clés

Open data, gouvernance des données, médiation, culture de l'information.

TITLE

Reinventions of democracy in the light of open data : from the discourse of participation to the constraints of governance

Abstract

Research on policies and practices of mediation, valuation and education on open data reveals that the opening of data triggers new questions that carry a fundamental question about

the place of data in the functioning of contemporary democracy. The public open data space is determined in negotiating locations and in unclear arrangements. The devices that allow access to these data respond more to a logic of administrative and economic governance than to democratic principles. Thus, for open data to participate in a form of reinvention of democracy in the context of *big data*, the information and cognitive practices that allow the use of data require complex acculturation actions and the development of data literacy.

Keywords

Open data, data governance, mediation, information literacy.

TÍTULO

Resumen

La investigación sobre políticas y prácticas de mediación, evaluación y educación en torno a los datos abiertos revela que es difícil encontrar un lugar en el dominio público de la democracia contemporánea. El espacio público de datos abiertos se determina en las ubicaciones de negociación y en los acuerdos poco claros. Los dispositivos que permiten el acceso a estos datos responden a una lógica de gobierno administrativo y económico que es más que democrática. Por lo tanto, para el propósito de la reinención de la democracia en el contexto del *big data*, la información y las prácticas cognitivas que permiten el uso de datos requieren acciones de aculturación que son complejas.

Palabras clave

Datos abiertos, gobierno de datos, mediación, cultura de la información.

INTRODUCTION

Parmi la masse des données qui alimentent le *big data*, les données ouvertes, open data, occupent une place particulière. Elles désignent les données collectées par les organismes publics ou privés chargés d'un service public et mises à disposition en format numérique sur des plateformes nationales ou locales permettant leur libre accès et leur réutilisation par les citoyens ou les organisations. Plusieurs types de données peuvent faire l'objet d'une ouverture et d'une mise à disposition pour la société (Liquète, Gilliard, 2017). Différents territoires sont concernés, dans la logique affirmée par la mission Etalab en 2011, confirmée par la création de la fonction d'administrateur général des données (décret n° 2014-1050 du 16 septembre 2014), la participation de la France au Partenariat pour un Gouvernement Ouvert en 2014 et la loi pour une République numérique en 2016. Cette dernière a ajouté la notion de "données d'intérêt général" qui peuvent être diffusées par les acteurs privés en lien avec des données publiques, comme ceux qui perçoivent des subventions publiques. Les données provenant de travaux de recherches scientifiques subventionnés par l'Etat en font partie. Un livre blanc sur les données ouvertes (Meszaros, Samath, Guerin-Hamdi, Faure, 2015) indique que les enjeux de ces dernières concernent les sphères socio-économiques, scientifiques et culturelles pour permettre une meilleure capacité intégrative de l'individu et du groupe social, des possibilités d'innovation, des potentialités de réaction et d'adaptation à l'évolution de l'organisation individuelle ou collective. Le portail de modernisation de l'action publique voit dans l'open data un levier du changement de l'administration, affirmant que « *l'open data nourrit la participation citoyenne et vice-versa, confère de nouveaux moyens d'agir et stimule la démocratie* »¹.

Mises à disposition du public, les données ouvertes découvrent un champ potentiel de connaissances et d'utilisations multiples, dans une perspective technologique de « *smart*

city », mais aussi de démocratie participative, à tous ceux qui sont en capacité de les appréhender et d'en faire usage. L'ouverture des données répond à un enjeu politique de transparence de l'action publique à travers une conception très large du droit à l'information des citoyens initiée en France par la loi du 17 juillet 1978 sur la liberté d'accès aux documents administratifs et de la réutilisation des informations publiques. Ce droit concerne le citoyen non seulement du point de vue de la protection de ses données personnelles (dans la logique confirmée par le Règlement général sur la protection des données qui entre en vigueur en mai 2018), mais également de sa possibilité de contrôler l'action publique, voire d'y participer directement. Ainsi, Sarah Labelle et Jean-Baptiste Le Corf (2012) ont montré que les politiques d'ouverture des données s'accompagnent de discours qui visent le politique à travers la qualité des processus démocratiques, l'administratif à travers une amélioration des relations entre administration et administrés, et le technique à travers le soutien à l'innovation. Dans la perspective de valoriser l'ensemble de ces objectifs, les enjeux de l'ouverture sont aussi communicationnels (Capelle, Lehmans, 2016). L'expression *Open data* est initialement liée à la recherche publique environnementale. 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 (Goncalves, Rufat, 2016). Concrètement, sur les portails d'*open data*, des jeux de données peuvent être téléchargés librement en vue d'un traitement qui permet, par exemple, de vérifier des hypothèses, de créer des applications, ou de représenter visuellement des informations grâce à des outils graphiques. Ils constituent pour le citoyen à la fois une fenêtre sur l'activité des services publics, une manne informationnelle alimentant le débat public, et un espace de participation tant par la réutilisation des données que par la contribution à leur collecte. Selon les jeux de données, la manière dont ils sont présentés, et l'exploitation qui est visée, les opérations techniques et cognitives nécessaires peuvent s'avérer complexes.

Si la question des données ouvertes évolue rapidement dans les agendas politiques, du local à l'international, la réalité des pratiques en France reste hétérogène (Goeta, 2016). L'affichage d'une politique d'ouverture cache les contraintes fortes qui pèsent sur la gouvernance des données. Il n'est pas exempt de discours idéologiques, comme le montre Etienne Damome (2018) à propos du cas africain, dans lequel le développement et la démocratisation sont au cœur d'un « fantasme techno-déterministe » qui accompagne la politique d'ouverture des données. Plus profondément, cette question interroge les fondements mêmes de la démocratie, quand la gestion des données risque d'envahir ou de remplacer l'espace public et d'en effacer les dimensions proprement politiques du débat, d'en faire oublier la nature contraignante. Une recherche sur les politiques et les pratiques de médiation, de valorisation et d'éducation autour des données ouvertes a été menée par notre équipe dans la région Nouvelle Aquitaine². L'enquête a porté sur plusieurs entités productrices et utilisatrices de données ouvertes dans l'académie de Bordeaux. Son objectif était d'analyser, d'une part, les politiques d'ouverture, de médiation et de médiatisation des données à l'échelon territorial, d'autre part, la façon dont les enseignants recourent effectivement à ces données ouvertes dans un objectif pédagogique. Elle révèle que l'ouverture des données déclenche des questions nouvelles qui peinent à trouver leur place dans l'espace public mais qui sont pourtant porteuses d'une interrogation fondamentale sur le fonctionnement de la démocratie contemporaine, à l'heure où l'extrême quantité, la variété et la vitesse de circulation des données qui caractérisent le *big data* semblent remettre en cause les processus traditionnels de prise de décision.

.....

1. <http://www.modernisation.gouv.fr/ladministration-change-avec-le-numerique/par-louverture-des-donnees-dans-les-administrations/open-data-nourrit-la-participation-citoyenne>

2. L'équipe RUDII (Représentations, usages, développements et ingénierie de l'information) de l'IMS (Intégration matériau-système, UMR5218), qui a mené le projet de recherche dans l'académie de Bordeaux. Le carnet de recherche Data-cultures est disponible : <http://dcultures.hypotheses.org/>

Alors que les pratiques journalistiques se déplacent sur la manipulation de données (*fact checking*, lutte contre les *fake news*) et de la *data visualisation*, que l'école s'ouvre de plus en plus aux techniques de gestion des données pour construire des connaissances, que les discours politiques s'appuient largement sur l'objectivité et la transparence apparentes des données et des chiffres, un temps de réflexion sur le sens et les usages des données ouvertes est indispensable. Quelle place sont-elles susceptibles d'occuper dans l'espace public entendu comme un espace de communication politique, du point de vue organisationnel, des représentations et des pratiques ? L'espace public de la donnée ouverte est un espace symbolique de discours articulé aux dispositifs socio-techniques de mise à disposition et d'usage des données (Lehmans, 2018). L'ouverture des données est associée dans les discours à la gouvernance ouverte, voire au gouvernement ou à la démocratie ouverte. Techniquement, les données ouvertes se présentent sous forme de bases contenant des jeux de données, constituées et mises à disposition dans le cadre d'une collecte réalisée par différents acteurs aux stratégies complémentaires, concurrentes ou divergentes. Elles n'existent pas tant qu'elles ne sont pas mises en visibilité, et socialement appropriées. La donnée est produite dans un contexte et offerte dans une perspective d'usage projeté. Il est donc important d'identifier les conditions de cette production d'une part, les projets d'usages d'autre part, les conditions de réalisation de ces projets enfin, pour les situer dans la perspective politique d'une évolution des mécanismes démocratiques. Dans un premier temps, il faut souligner que l'espace public des données est déterminé dans des lieux de négociation et selon des arrangements qui doivent être décryptés. Ensuite, l'analyse des dispositifs qui permettent l'accès aux données montre qu'ils répondent à une logique de formatage organisationnel et technique à considérer en fonction des préoccupations de participation. Enfin, les pratiques informationnelles et cognitives qui permettent les usages des données nécessitent des formes d'acculturation qui sont complexes et peu négociables, et pourtant conditions sine qua non de l'occupation maîtrisée d'un espace public de la donnée transformable en lieu de décision et de participation politique démocratique.

GOUVERNANCE OUVERTE ET LISIBILITÉ DE L'ESPACE PUBLIC DE LA DONNÉE

Les données ouvertes relèvent d'une politique volontariste de partage et de collaboration dans la production et les usages de l'information à partir de données mises à disposition. Du point de vue des politiques publiques, elles reposent sur la mise en œuvre de principes affichés 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). Le premier principe de transparence se heurte aux contraintes organisationnelles des entités productrices de données.

Des stratégies politiques de mise en visibilité des données

Avant la loi pour une République numérique, qui pose le principe de l'ouverture des données publiques par défaut, la décision d'ouverture et de diffusion des données est souvent politique et parfois liée à une concurrence entre collectivités sur un territoire ou entre services dans une collectivité. Elle révèle des stratégies de communication politique autour de la transparence et de l'innovation qui peuvent être mises en avant par des collectivités de taille très modeste comme la commune de Brocas Les Forges dans les Landes (800 habitants), qui affiche une politique volontariste d'ouverture des données depuis 2011 et un site avec des propositions de réutilisations des données. Ainsi, l'ouverture des données s'inscrit dans une rhétorique de la visibilité, de l'innovation et de la participation, parfois autour de la thématique plus précise de l'économie sociale et solidaire. La conséquence de ces stratégies est que les données sont souvent orientées vers les

informations géographiques et liées à des territoires, et que chaque territoire développe ses propres pratiques d'ouverture. Des phénomènes d'asymétries d'information entre acteurs et territoires (Chartron, Broudoux, 2015) risquent d'entrer en contradiction avec la logique d'ouverture des données qui, pour être efficace à grande échelle et produire des effets durables, ne peut pas être cloisonnée. En outre, les logiques communicationnelles d'ouverture ne sont pas toujours congruentes avec les logiques organisationnelles.

Le projet de recherche a permis d'enquêter entre 2016 et 2017 sur le terrain de collectivités locales productrices de données auprès d'acteurs en Nouvelle Aquitaine, des médiateurs et des utilisateurs finaux des jeux de données ouvertes. Du côté des producteurs de données, personnels de collectivités, médiateurs, élus, des observations participatives de séances de travail sous forme d'ateliers et de focus groupes ont été réalisées, ainsi que des entretiens semi-directifs auprès de six agents, afin d'identifier et de caractériser le discours des acteurs de terrain pour comprendre les objectifs qu'ils se fixent, et leur représentation de la chaîne de production et de diffusion des données dans la sphère publique actuelle. Les entretiens permettent de comparer les représentations d'acteurs entre des secteurs d'activités ou des échelles géographiques différents, de comprendre l'articulation de la chaîne de production et de diffusion des données, et de saisir l'évolution de leurs pratiques déclarées sur leurs terrains. Du côté des utilisateurs dans un objectif d'éducation, l'équipe de recherche s'est intéressée aux usages des enseignants. Elle l'a fait à partir d'entretiens menés auprès de quinze d'entre eux, intervenant à l'université, dans la formation des professeurs, ou au rectorat dans la formation continue, et d'observations de séances de formation d'étudiants ou d'enseignants. Tous les enseignants interrogés étaient des utilisateurs potentiels ou avérés de données ouvertes dans leurs pratiques pédagogiques.

Plusieurs plateformes d'*open data* ont vu le jour au sein des collectivités publiques pour présenter et mettre à disposition leurs jeux de données. En France, Etalab a mis en place la plateforme data.gouv.fr, tandis que des collectivités territoriales ont créé leurs propres plateformes de données locales et que des entreprises comme la SNCF, mais aussi Uber ou AirBnB, communiquent sur l'ouverture de (certaines de) leurs données. Dans la Région Nouvelle Aquitaine, les efforts se sont multipliés pour mettre à disposition des données dans le domaine du tourisme (SIRTAQUI), de l'environnement et des données géographiques (PIGMA, Datalocale), par exemple. Ces plateformes numériques reposent sur des stratégies de mise en visibilité des données (Mabi, 2015) par une grande variété d'acteurs qui sont entrés dans des processus d'ouverture pour des raisons diverses : stratégiques et économiques, lorsque que le calcul coût-bénéfice penche en faveur de l'ouverture pour des plateformes d'information géographique par exemple, politiques, lorsque des élus s'emparent de la question des données ouvertes dans une perspective militante ou d'image, normatives, lorsque la loi impose l'ouverture des données. Dans les collectivités territoriales, la direction numérique, à partir d'une demande politique ou de sa propre initiative, donne souvent l'impulsion aux propositions qui sont validées suivant un schéma administratif puis par les élus qui interviennent en fin de processus. Ces derniers n'ont pas toujours une compréhension globale et précise des enjeux de l'ouverture des données.

DES CONTRAINTES ORGANISATIONNELLES DE FAIBLE VISIBILITÉ DES DONNÉES

Malgré la législation récente, les administrations sont encore loin de mettre à disposition toutes les données qu'elles gèrent. Les agents des collectivités territoriales témoignent d'une grande diversité des pratiques entre les services, qui ne sont pas tous, du point de vue technique, politique, ou managérial, prêts à considérer cette gestion dans un écosystème ouvert. L'ouverture des données implique la transformation des structures et des

modes de gouvernance des organisations ainsi que des compétences des agents. Dans le rapport remis par le Secrétariat général pour la modernisation de l'action publique en 2015, la gouvernance de la donnée désigne « *l'ensemble de principes et de pratiques qui visent à assurer la meilleure exploitation du potentiel des données* » (p. 48). Les principes relèvent d'une logique de transversalité dans l'organisation, les pratiques d'une dynamique de communication. 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 une évolution des systèmes d'information et un décloisonnement des administrations. La gouvernance « *associe négociations, prises de décision et émergence de consensus* » (Schafer et Le Crosnier, 2011). La gouvernance ouverte lie les principes de transparence et de participation dans une relation complexe intégrant visibilité de l'action publique, développement des moyens d'expression, voire d'action, des citoyens et "comportements organisationnels" ou logiques bureaucratiques qui peuvent constituer des entraves à l'ouverture (Pasquier, Villeneuve, 2007).

L'ouverture des données dans une organisation questionne ainsi ses modes de fonctionnement interne, les compétences professionnelles ou « métiers » dans et entre services, la façon d'intégrer le public susceptible d'utiliser les données dans la prise de décision, et, de façon générale, l'écosystème informationnel et les interactions entre les services et les acteurs. Les méthodes de collecte, de traitement, de conservation et de communication des données sont propres à chaque service, qui fonctionne avec la direction des systèmes d'information, mais pas de façon transversale et partagée avec les autres services. Les problématiques d'accessibilité et d'utilisabilité perturbent les fonctionnements routiniers de services cloisonnés dans leur gestion des données, 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, porteurs d'un discours d'ouverture et de transparence. Le dossier de l'ouverture des données est impulsé par un service en particulier, souvent la direction des systèmes d'information qui en fait un levier de modernisation des pratiques, mais parfois aussi les services nouveaux créés autour des questions d'usages, portés par des agents dont les compétences sortent du seul domaine de l'informatique et sont en lien avec la gestion de l'information d'une part, l'intelligence économique d'autre part, l'animation des données enfin. Ainsi, pour l'un des acteurs interrogés, « *dans une organisation hyper-hiérarchique, soit on sensibilise les cadres, soit on fonctionne en mode projet. On essaie de ne pas faire de l'open data une fin en soi, sinon on recrée un silo.* »

De nombreux métiers nouveaux se sont développés autour des stratégies d'ouverture des données, qui n'ont pas encore totalement acquis leur identité dans la langue française puisque l'on continue de parler de « *data scientist* », « *data analyst* », « *data manager* », « *chief data officer* » ou « *data protection officer* », cette dernière fonction bientôt rendue obligatoire par l'application du Règlement Général sur la Protection des Données européen. La question de la donnée dans le *big data* est un enjeu de pouvoir central dans le fonctionnement des organisations aujourd'hui, et la problématique des données ouvertes renforce encore la position des professionnels de l'informatique, tout en nécessitant des formes de coopération et de communication avec les professionnels de l'information qui ont une vision des enjeux économiques et sociaux. La culture de la donnée n'est pas strictement technique mais aussi analytique, liée à la capacité d'identifier les enjeux et les données susceptibles de créer de l'information qui a de la valeur. La gouvernance reste élitiste dans ses principes de fonctionnement, même si ceux qui portent les projets d'ouverture et de mise à disposition des données se vivent souvent comme des missionnaires. Ainsi, dans une collectivité, c'est un « conseil de sachants » qui donne l'impulsion dans la prise de décisions, parce que ceux qui savent connaissent les contraintes et sont conscients des

enjeux. Les enjeux politiques de l'ouverture des données sont fortement corrélés à des questions économiques dans le cadre global du *big data*.

FORMATAGE ET UTILISABILITÉ DE L'ESPACE PUBLIC DE LA DONNÉE

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, au cœur des principes de gouvernance de l'information dans les entreprises, n'apparaît pas de façon évidente dans les discours des acteurs 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 autour de la thématique de la participation, l'autre sur l'expertise technique et l'efficacité autour de la thématique de la modernisation et de l'innovation. C'est clairement l'entrée des usagers et la question de l'utilisabilité des données dans l'écosystème qui imposent la nécessité de repenser les formats d'interopérabilité sur le plan technique.

Repenser les formats

Avant la mise en place d'une dynamique d'ouverture, les données étaient collectées et les flux traités et régulés dans les systèmes internes d'information sans que les questions de l'ouverture, de la publication et finalement de la communication vers l'extérieur fussent prises en compte. L'ouverture des données réinterroge le cycle de vie de la donnée en contraignant les organisations et les services à prendre en considération, dès le moment de la collecte, la possibilité de réutilisation, sans que celle-ci soit clairement définie à l'avance. En effet, la mise à disposition des données n'implique pas automatiquement leur usage. Le volume, la diversité et l'hétérogénéité des données nécessitent un pilotage dans le traitement de l'infrastructure technique et logicielle ainsi que dans la description et la mise en forme des données en vue de leur publication, dans une logique de formatage comme condition du partage. Le pilotage doit intégrer la valeur des données au regard des usages projetés, de la transformation de la donnée en information, dans une démarche de qualification qui rompt avec les routines organisationnelles et qui nécessite de la part des agents une expertise dans la représentation des bases de données à l'aune de leurs usages externes. L'utilisabilité des données doit ainsi être considérée, définie par la norme comme « *le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficacité et satisfaction, dans un contexte d'utilisation spécifié* » (ISO 9241-11:2018). Le format informatique, la licence, l'organisation de la base de données, conditionnent l'utilisabilité et la réutilisabilité des données ouvertes et les installent dans une logique de normalisation.

Les collectivités ont commencé à mettre en place des formes de réorganisation des procédures et des métiers quand elles se sont lancées volontairement dans des politiques d'ouverture. Elles sont désormais contraintes de le faire avec la mise en application du Règlement général pour la protection des données, qui impose le principe de responsabilité par rapport aux usages. Outre la gestion de la protection des données personnelles, une dimension incontournable dans cette réorganisation est celle de la normalisation qui porte des enjeux juridiques et techniques mais aussi fonctionnels. Les normes et les formats assurent le lien entre collecte, accès et usages tout au long du cycle de vie de la

donnée autour duquel les compétences des agents se réforment. Dans un contexte de territorialisation des données ouvertes et de coexistence entre des pratiques de gestion des données diverses, l'absence de normalisation freine l'interopérabilité, la garantie de la qualité et la possibilité de diffuser des jeux de données universellement utilisables.

Standardisation et usages de l'information

La standardisation peut constituer un levier de gouvernance des données, quand elle devient un objet partagé dans et entre les services. Dans le domaine de l'information géographique, la directive INSPIRE de la Direction générale de l'environnement de la Commission européenne datant de 2007, « *visé 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 à des plateformes comme PIGMA de fonctionner. L'application de cette directive oblige, dans les organisations chargées de l'information géographique, les thématiciens et les informaticiens à adopter des normes, des langages, des pratiques communs visant l'interopérabilité des données. Une véritable culture technique et normative s'est développée sur la base de cette directive pourtant très complexe, et a donné aux services d'information géographique une longueur d'avance sur les collectivités peu sensibilisées aux nécessités de standardisation et à des outils techniquement complexes. 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.

Dans une logique de partage de la donnée, les formats conditionnent l'exploitabilité des données. L'utilisation d'un format ouvert (de type .csv) pour mettre à disposition les jeux de données est essentielle. Les licences qui déterminent les conditions juridiques de réutilisation des données sont également fondamentales. Elles devraient privilégier le partage à l'identique, mais celui-ci est contraire aux intérêts commerciaux des entreprises et un frein à l'utilisation des données ouvertes. La plupart des collectivités territoriales et l'association Open Data France ont donc renoncé à imposer des licences trop strictes (OdB) 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). Enfin, la qualité des métadonnées est très importante pour permettre la contextualisation des jeux de données. Charlotte Maday (2015 : 161) souligne ainsi la proximité de la gestion des données ouvertes avec le *record management*.

L'utilisation réelle et contextualisée des données requiert le respect de normes et de standards dont les exigences techniques ne sont pas toujours aisées à comprendre, notamment des élus, et qui conditionnent pourtant l'efficacité des politiques d'ouverture. Mais celles-ci, au niveau territorial, doivent avant tout rencontrer des besoins et des usages, une culture partagée de la donnée.

CULTURES DE LA DONNÉE, DÉMOCRATISATION DE L'ESPACE PUBLIC DE LA DONNÉE ET COMMUNS DE LA CONNAISSANCE

L'ouverture reste une utopie (Goeta, 2016) si les conditions organisationnelles et culturelles de leur usage ne sont pas réunies. Même si les données sont accessibles, sur les plans technique et juridique, le « public » visé par l'ouverture doit connaître leur existence

d'une part, être capable de les utiliser d'autre part. Les responsables des services chargés de l'ouverture des données le savent, le public ciblé reste nébuleux. Les entreprises, ou les services publics qui ont besoin de données, sont actuellement les principaux bénéficiaires de l'ouverture des données. Les conditions d'une véritable réinvention ou rénovation de la démocratie à l'aune de la participation de tous aux usages des données, dans le cadre de communs informationnels (Peugeot, 2016), restent à construire, par des stratégies de médiation et d'éducation. Ce qui pourrait constituer la part démocratique du *big data* à travers la création d'un espace commun des données n'est pas encore d'actualité. Cet objectif repose sur deux exigences de base qui sont prises en charge par les organisations : la première est que les données soient visibles et lisibles par tous grâce à des bases de données permettant de les rechercher aisément ; la seconde est qu'elles soient suffisamment accessibles, pour être utilisées par des acteurs qui n'ont pas participé à leur processus de fabrication. Une troisième dimension reste essentielle, mais elle est encore peu envisagée : le développement d'une littératie des données ouvertes à travers des stratégies de médiation et d'éducation. C'est cette dimension qui a été analysée dans le projet de recherche.

Les stratégies de médiation

Les objectifs de soutien à l'innovation avec l'open data énoncés dans la loi Pour une République Numérique (2016), et l'ouverture d'une manne informationnelle en direction de la société civile, peuvent favoriser de nouveaux usages à partir des données. Mais l'innovation concerne essentiellement, pour l'heure, les moyennes ou grandes entreprises qui disposent de ressources humaines et techniques suffisantes pour valoriser les données, à la différence des petites entreprises et des citoyens. L'usage des données, dans le monde du *big data* et de l'information, est complexe et coûteux, tant il suppose l'intervention de spécialistes. En outre, les données ne sont utilisables que dans le cadre d'un travail complexe de traitement et de construction de l'information à la base des connaissances. Il apparaît alors essentiel de développer des offres de services et de formations qui permettent aux citoyens et aux entreprises d'utiliser les données ouvertes telles qu'elles sont produites et mises à disposition, en vue de répondre à des besoins d'information pour agir, dans une logique d'innovation et d'économie sociale et solidaire.

Quels que soient les moyens de médiation mis en place sur les plateformes, la manipulation des données par traitement informatique à l'aide de machines suppose la maîtrise de compétences de la part des utilisateurs. Sur la plateforme data.gouv.fr, les "meilleures" réutilisations sont rendues visibles et mises en avant sur la page d'accueil, ce qui valorise les usages possibles et invite les contributeurs potentiels à la créativité. Des explications sur les différentes méthodes, moyens et outils techniques accompagnent aussi les usagers qui souhaitent se lancer en tant que contributeurs, mais de façon très sommaire. Dans les collectivités territoriales, des actions de médiation sont mises en place en interne pour diffuser la culture des données ouvertes dans et entre les services, souvent avec l'aide de médiateurs extérieurs comme la Fondation internet nouvelle génération, qui organise des « Infolabs ». En direction de l'extérieur et des publics, outre la médiation numérique *via* l'organisation et la communication autour des plateformes, ce sont plutôt les « animations » qui sont privilégiées ; cartoparties, ateliers de réflexion et de data-visualisation, concours, appels à projets permettent de mettre en scène les données ouvertes en direction de publics précisément ciblés et dont la détermination est liée aux missions et au ciblage politique de la collectivité.

Cependant, les seules visibilité en ligne et animations des données ouvertes ne permettent pas aujourd'hui de répondre aux promesses de l'*open data*, en matière de réutilisation des données par tous, d'intelligence économique et plus généralement d'innovation et de participation. Le développement des compétences liées au numérique et plus précisé-

ment aux données est indispensable, comme le rappelle la création par un décret du 3 novembre 2017 du haut-commissaire à la transformation des compétences, auprès du ministre du travail. Cette nécessité était affirmée par un avis du Conseil économique, social et environnemental dès 2015 (Pérès, 2015). Pour les diffuseurs de données, l'éducation est une frontière indépassable.

Les stratégies d'éducation

Même si l'idée que le « grand public » puisse aller puiser dans les banques de données reste utopique, l'importance de diffuser une *data literacy*, une culture qui permette à chacun de lire, créer et communiquer des données, tout au moins d'en comprendre les enjeux, semble essentielle. L'un des objectifs du projet de recherche était d'étudier, en regard des pratiques déclarées des diffuseurs de données, les pratiques d'enseignants et de formateurs, de l'école primaire à l'université, recourant aux données ouvertes dans un objectif pédagogique avec leurs élèves et leurs étudiants. On constate que peu d'enseignants se sont encore saisis de ces ressources. Ceux qui l'ont fait ont mobilisé beaucoup d'énergie et ont immédiatement trouvé des réseaux d'échange (Travaux académiques mutualisés dédiés aux données ouvertes dans le domaine de l'éducation aux médias en 2017-2017 et des sciences économiques en 2017-2018, notamment). Bien qu'intéressés par le sujet, ils se confrontent à sa complexité. Certains éléments d'enquête, issus des entretiens individuels auprès d'enseignants, démontrent l'existence de réelles difficultés à identifier et définir les données ouvertes ainsi qu'à envisager leurs usages en situation d'enseignement-apprentissage.

Les données ouvertes constituent cependant un enjeu éducatif central autour du *big data* en vue de la maîtrise par les futurs citoyens des données qui les concernent personnellement, dans leurs activités ou dans leurs prises de décision, y compris sur le plan politique. Plusieurs compétences sont visées, permettant d'acquérir une culture analytique et critique des éléments de traitement, de communication et de réutilisation des données. Cette culture intègre des compétences complexes comme la capacité à trouver, nettoyer et exploiter des jeux de données, à produire des statistiques, à traduire des questions liées aux connaissances scolaires dans une datavisualisation (Capelle, Lehmans, Liquète, 2017), jusqu'à la compréhension des « algorithmes qui transforment et traduisent les données en visualisation » (Desfriches Doria, 2015, p. 54). Pour les enseignants, faire avec cette complexité des techniques, mais aussi des enjeux, demande beaucoup de temps, souvent trop par rapport à celui dont ils disposent pour former des « futurs citoyens » dans le cadre de leurs programmes. Les stratégies pédagogiques les plus élaborées sont basées sur la création de données nouvelles, dans une perspective de participation (Capelle, Lehmans, Liquète, 2017). Elles consistent à faire produire par les élèves eux-mêmes des données, en vue de la construction d'information puis de connaissance, en cartographiant par exemple des éléments du patrimoine, ou des données utiles aux déplacements des personnes en situation de handicap. La culture des données peut ainsi constituer une partie des « entourages qui procurent aux individus les conditions de leur liberté » (Zask, 2011, 58) dans un espace public où la participation relève d'une expérience concrète. Participer, comme le souligne Joëlle Zask avec John Dewey (2003), c'est prendre part à une enquête qui contribue à fabriquer le commun.

CONCLUSION

Les données sont un élément de base du dialogue ouvert qui se noue entre les collectivités publiques, les entreprises et les citoyens. Ces derniers sont des acteurs essentiels dans le

processus d'ouverture des données, en tant qu'utilisateurs mais aussi acteurs potentiels d'une collaboration pour identifier les besoins ou interroger les processus, les espaces et les méthodes pour des usages sociaux et économiques diversifiés, dans un processus de co-construction ou de débats. Mais pour qu'ils puissent réaliser ce potentiel, un effort d'éducation et de formation reste à faire. Quelques enseignants en sont conscients, tout en soulignant la complexité d'une question qui tisse des enjeux politiques et techniques. Nous faisons le postulat que la valorisation des données ouvertes peut constituer un élément de "communs de la connaissance" à construire (Latour, 2011). L'analyse des processus d'ouverture montre que la mise à disposition de données relève de stratégies variées et parfois contradictoires. Dans les représentations des élus et des cadres, les enjeux sont essentiellement communicationnels, lorsqu'il s'agit de ne pas rester en dehors de la force d'attraction du *big data*. Ils peuvent également être organisationnels, lorsqu'il s'agit de rénover le fonctionnement des services, de décroiser les circuits fermés de l'information et de réviser leur articulation au regard du système d'information. Ils peuvent enfin relever d'un véritable projet politique de développement d'espaces publics de la donnée, de l'information et de la connaissance. L'espace public des données ouvertes est un espace à conquérir et à occuper pour devenir commun.

RÉFÉRENCES BIBLIOGRAPHIQUES

Capelle, Camille, Lehmans, Anne, Liqueste, Vincent (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. URL : halshs01633284.

Chartron, Ghislaine, Broudoux, Évelyne (2015), « Enjeux géopolitiques des données, asymétries déterminantes », in *Big Data - Open Data : Quelles valeurs ? Quels enjeux ?* Actes du colloque « Document numérique et société », Rabat, De Boeck Supérieur, p. 65-83.

Damome, Étienne (2018), « Opportunités et difficultés du développement des archives ouvertes pour la communication publique : la situation en Afrique subsaharienne », *Revue française des sciences de l'information et de la communication* [En ligne], <http://journals.openedition.org/rfsic/3491>.

Desfriches Doria, Orélie (2015). « Quels dispositifs numériques pour appréhender la data-visualisation ? », *I2D – Information, données & documents*, vol. 52, n° 2, p. 54-56.

Dewey, John (2003), *Le public et ses problèmes*, trad. et introd. J. Zask, Pau : Farrago / Léo Scheer.

El Hachani, Mabrouka (2015), « Open data, collectivités et usagers : une dynamique en question », in Paquienséguy, Françoise (dir.), *Open data. Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, Paris : Éditions des archives contemporaines, p. 1-23.

FRANCE. Secrétariat général pour la modernisation de l'action publique (2016), *Rapport d'activité 2015*, <http://www.ladocumentationfrancaise.fr/rapports-publics/164000176/index.shtml>

Goeta Samuel (2016), « Instaurer des données, instaurer des publics : une enquête sociologique dans les coulisses de l'open data ». Thèse en sociologie, Télécom ParisTech, URL : (tel-01458098)

Goncalves, Dann, Rufat, Samuel (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://journals.openedition.org/cybergeog/27750>.

Labelle, Sarah, Le Corf, Jean-Baptiste (2012), « Modalités de diffusion et processus documentaires, conditions du « détachement » des informations publiques. Analyse des discours législatifs et des portails open data territoriaux », in Bardou Boissier Sylvie et Pailliarth Isabelle (coord.), Dossier Information publique : stratégies de production, dispositifs de diffusion et usages sociaux, *Les Enjeux de l'Information et de la Communication*, n° 13-2, p. 59-71, URL : <https://lesenjeux.univ-grenoble-alpes.fr/2012/dossier/05-modalites-de-diffusion-et-processus-documentaires-conditions-du-detachement-des-informations-publiques-analyse-des-discours-legislatifs-et-des-portails-open-data-territoriaux>

Latour, Bruno (2011), « Il n'y a pas de monde commun : il faut le composer », *Multitudes*, vol. 45, n° 2, p. 38-41.

Lehmans, Anne, Capelle, Camille (2016), « Gouvernance et horizon d'attente des données ouvertes pour l'éducation », *Actes du 12ème colloque EUTIC*, Zakyntos, Grèce, 2016, URL : (hal-01790492).

Lehmans, Anne (2018), « L'horizon d'une culture de la donnée ouverte : de l'utopie aux pratiques de gouvernance des données », *Revue COSSI*, n° 1, URL : <http://www.revue-cossi.info/numeros/n-1-2018-big-data-thick-data/708-1-2018-revue-lehmans#citer>

Liquète, Vincent, Gilliard, Armelle (2017), « Épistémologie de la donnée au risque de la connaissance : approches croisées. », Communication au colloque COSSI *Méthodes et stratégies de gestion de l'information par les organisations : des «big data» aux «thick data»*, 11 et 12 mai 2017, Université McGill, Montréal, Canada.

Mabi, Clément (2015), « La plate-forme « data.gouv.fr » ou l'open data à la française », *Informations sociales*, vol. 191, n° 5, p. 52-59.

Maday, Charlotte (2015), « L'apport de la gestion des documents d'activité (records management) à l'ouverture des données. Réflexions basées sur les pratiques en France », *Les cahiers du numérique*, vol. 11, n° 2, p. 149-166.

Maurel Dominique, Chebbi Aïda (2012), « La perception de la confiance informationnelle. Impacts sur les comportements informationnels et les pratiques documentaires en contexte organisationnel », *Communication & Organisation*, vol. 2, n° 42, p. 73-90.

Meszaros, Branislav, Samath, Sitthida, Guérin-Hamdi, Sonia, Faure, Céline (2015), *Livre blanc sur les données ouvertes*, URL : <halshs-01162692>.

Pasquier, Martial, Villeneuve, Jean-Patrick (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*, vol. 73, n° 1, p. 163-179.

Pérès, Éric (2015), « Les données numériques : un enjeu d'éducation et de citoyenneté », Avis du Conseil économique, social et environnemental, URL : http://www.lecese.fr/sites/default/files/pdf/Avis/2015/2015_01_donnees_numeriques.pdf

Peugeot, Valérie (2016), « Facilitatrice, protectrice, institutrice, contributrice : la loi et les communs », Contribution au colloque de Cerisy - *Vers une république des biens communs ?*, Septembre 2016, URL : <https://vecam.org/Facilitatrice-protectrice-institutrice-contributrice-la-loi-et-les>

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

Zask, Joëlle (2011), *Participer. Essai sur les formes démocratiques de la participation*, Lormont : Le Bord de l'eau.