



HAL
open science

Gouvernance et horizon d'attente des données ouvertes pour l'éducation

Anne Lehmans, Camille Capelle

► **To cite this version:**

Anne Lehmans, Camille Capelle. Gouvernance et horizon d'attente des données ouvertes pour l'éducation. XIIe Colloque International EUTIC 2016 Réseau de recherche européen et pluridisciplinaire sur les enjeux et usages des TIC LOGIQUES DE RÉSEAUX ET NOUVELLES GOUVERNANCES Université Nationale et Kapodistrienne d'Athènes , Sep 2016, Zakynthos, Grèce. hal-01790492

HAL Id: hal-01790492

<https://hal.science/hal-01790492v1>

Submitted on 12 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gouvernance et horizon d'attente des données ouvertes pour l'éducation

Anne LEHMANS, Camille CAPELLE

Laboratoire IMS UMR 5218 CNRS, Université de Bordeaux, France

anne.lehmans@u-bordeaux.fr

camille.capelle@u-bordeaux.fr

Résumé

La mise à disposition, l'accessibilité publique et l'utilisation des données ouvertes s'inscrivent dans un projet de partage d'informations et dans une logique de communs de la connaissance qui ont introduit une possibilité de gouvernance ouverte particulière et un nouvel enjeu de société pour les citoyens, dont l'école peut se saisir. Pourtant, les données ouvertes restent en dehors de l'écosystème informationnel de la plupart des structures de formation. La présente recherche permet de définir les enjeux de la gouvernance des données ouvertes pour l'éducation à partir d'un état des lieux de la question et d'une première enquête qui met en évidence les discours et les moyens développés par les acteurs sur les modes de gouvernance susceptibles de fédérer une culture du partage des données, de permettre de scénariser ces données à l'intérieur de programmes de formation, de participer à la construction d'une culture de l'information autour de la "data literacy".

Mots-clés : données ouvertes, culture de l'information, gouvernance de l'information, innovation, éducation

Abstract

The provision, public accessibility and use of open data are part of an information-sharing project in a philosophy of knowledge commons that introduced a particular form of open government and a new social and educational issue for citizen. Yet open data remain outside the academic information ecosystem. This research defines the issues of open data governance for education from an inquiry that highlights the actors' discourses and arguments aiming to create a culture of data sharing, data display within training programs, and to build a "data literacy".

Keywords : open data, information literacy, information governance, innovation, education.

Introduction

Dans le domaine des sciences de l'information et de la communication, les chercheurs travaillent depuis plusieurs années sur les problématiques d'ouverture et d'exploitation des données publiques. Cette diffusion des données publiques est souvent présentée comme faisant partie intégrante de la mission de service public (Labelle, Le Corf, 2012). Les "données ouvertes" (*open data*) désignent des chiffres, relevés, mesures, réponses à des enquêtes, statistiques, comptages, et autres données quantitatives collectées par les organismes publics ou privés et mises à disposition en format numérique sur des plateformes nationales ou locales permettant leur accès et leur réutilisation par les citoyens ou les entreprises. Essentiellement quantitatives, elles peuvent être traitées, triées, croisées, exploitées, et visualisées à des fins de recherche, de développements ou d'enseignement. Elles ont des caractéristiques spécifiques dans l'économie de l'information, puisqu'elles sont variées, gratuites (même si elles ont un coût de production pour la collectivité), ouvertes et exploitables, mais ne peuvent avoir de sens sans traitements informationnels permettant leur intégration aux activités des professionnels ou du public (analyse, évaluation, visualisation, cartographie, etc.). Si la transparence et la participation sont au coeur du discours politique sur l'ouverture des données, leur utilisation correspond à des stratégies informationnelles spécifiques. Leur mise à disposition et leur accessibilité publique s'inscrivent dans un projet d'ouverture et de partage d'informations et dans une logique de "communs de la connaissance" (Hess et Ostrom, 2007) qui a introduit une possibilité de "gouvernance" ouverte particulière et un nouvel enjeu de société pour les citoyens, dont l'école peut se saisir. Leur traitement nécessite des compétences informationnelles, documentaires mais aussi techniques (informatiques) et soulève les questions socio-économiques, politiques ou géographiques. Manipuler ce type de données peut ainsi permettre de découvrir et de développer une culture numérique telle qu'elle est présentée par exemple dans l'Education aux Médias et à l'Information dès le collège en cycle 3 ou dans le programme d'enseignement au lycée en Informatique et Création Numérique. Pourtant, jusqu'à présent, les données ouvertes restent encore largement en dehors de l'écosystème informationnel de la plupart des structures de formation. Les enseignants n'ont pas toujours les moyens d'en prendre connaissance, ni de se former ou de les utiliser, encore moins de les organiser dans une dynamique de formation pour valoriser collectivement les transferts de savoirs. Faire entrer ces données dans l'enseignement nécessite l'implication d'une chaîne d'acteurs qu'il faut mobiliser dans une dynamique collaborative.

Cette question est à considérer dans le cadre plus large de l'émergence, dans le champ des techniques, des savoirs, et des imaginaires, de la donnée comme élément central des stratégies d'information et de connaissance. Cette émergence correspond à un mouvement d'externalisation de la connaissance vers l'environnement (Cardon 2014) qui interpelle directement l'école à plusieurs titres. La pyramide de la donnée à l'information à la connaissance et à

la sagesse (Ackoff, 1989), transformée en chaîne de la donnée à la compétence dans le Glossaire de la donnée publique (2016), sont des métaphores très claires de la place centrale de la donnée comme élément de base : de l'information en général, des informations personnelles, de la capacité à agir sur le monde en produisant et en utilisant des données de façon éclairée. Le développement de la thématique du "big data" dans le monde économique et social, de l'apprentissage du "code" dans le monde éducatif, montrent que la construction de l'autonomie cognitive et citoyenne est supposée passer aujourd'hui par la capacité à identifier, analyser et manipuler les données dans des systèmes informatiques. Du point de vue des politiques publiques, *l'open data* repose sur les trois principes de la transparence, la participation citoyenne et la 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), dans le cadre du web de données, qui s'installe irréversiblement comme le modèle d'avenir du web depuis l'annonce de son inventeur, Tim Berners Lee, dès 1994 (Kembellec et Prime-Claverie 2016 : 29). Ces problématiques nous incitent à nous interroger sur la possibilité de création de *communautés épistémiques* (Proulx, 2015) qui construisent de façon ouverte et collaborative, voire participative, des connaissances dans des domaines scientifiques ou scolaires, et peuvent donner corps à des modèles pédagogiques et citoyens innovants. Le mouvement de libération des données repose sur la recherche de la désintermédiation et de la «démocratisation» de l'information, mais le souci réaliste d'un usage par le public dans sa diversité impose des formes de médiation ou d'acculturation informationnelle, qui passent en particulier par l'éducation. 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.

Cette contribution propose une réflexion sur les enjeux de la gouvernance des données ouvertes pour l'éducation, à partir d'un état des lieux de la question et d'une première enquête menée dans une approche socio-pragmatique et compréhensive, sur la base d'entretiens et d'une analyse de discours et de pratiques, auprès de militants de données ouvertes, tant du côté des producteurs que de celui des récepteurs (institutions d'enseignement, chercheurs, enseignants). Cette proposition se situe en amont d'une recherche qui débute autour des données pour l'éducation. Elle vise plutôt la construction d'un cadre conceptuel et l'identification des enjeux d'une question informationnelle fondamentale. Cette première étude de terrain permet d'identifier et de cerner la démarche de différents acteurs qui réfléchissent, seuls ou collectivement, pour faire en sorte que les formes de mise à disposition des données, leur éditorialisation, la médiation qui les accompagne, puissent donner naissance à de nouvelles pratiques sociales et informationnelles, notamment celle d'apprendre dans la société du numérique. Nous cherchons à comprendre quels modes de gouvernance des données sont envisageables et envisagés par les différents acteurs, à quel moment, et selon quelles modalités. Après avoir précisé la question du lien entre gouvernance des données ouvertes et éducation, nous présenterons nos choix

méthodologiques, puis les premiers éléments qui nous permettent d'identifier des enjeux essentiels.

Démocratisation de l'accès aux données ouvertes : vers une éducation à la culture des données ?

En quête d'un mode de gouvernance innovant des données

L'ouverture lente mais progressive des données publiques se manifeste depuis les dernières années par leur accessibilité et la communication des institutions via des sites internet publics dédiés. Ces données, rendues accessibles et exploitables, sont livrées sans réelle concertation avec le public, parfois essentiellement dans une perspective de communication politique, mais dans l'attente affichée d'une appropriation de sa part. Considérant ces dimensions politiques et symboliques de l'ouverture du web et des données aujourd'hui, les pouvoirs publics, influencés par un mouvement militant et une prise de conscience citoyenne, seraient-ils en quête d'un mode de gouvernance des données innovant, à travers la mobilisation, l'engagement et la participation citoyenne ? La gouvernance, dans le contexte du web, suppose qu'on « *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 l'usage des données. Le concept de gouvernance ouverte lie transparence et 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 de la spécificité des publics dans leurs usages de l'information, mais également des "comportements organisationnels" qui peuvent constituer des entraves à l'ouverture (Pasquier, Villeneuve, 2007). Enfin, l'ouverture 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). Les enjeux politiques de l'ouverture des données sont fortement liés à des questions économiques dans le cadre global du *big data*. Au-delà de la prise en considération de ces 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 coeur des principes de gouvernance de l'information dans les entreprises, n'apparaît pas de façon évidente dans les discours et dans les normes concernant les données ouvertes.

La disparition de la médiation et l'accès direct aux données pour les usagers citoyens ne vont donc pas de soi. Si le mouvement de libération des données repose sur la recherche de la désintermédiation et de la « démocratisation » de l'information, la nécessité d'une forme de médiation et/ou d'une acculturation informationnelle demeure, en particulier dans l'éducation. Dans les recherches actuelles, l'intérêt des données ouvertes dans l'éducation est pour l'instant perçu en termes d'optimisation du système éducatif (Ben Adballah, 2016 : 42) ou à travers l'émergence d'une littératie nouvelle (*data literacy*) permettant de rapprocher les citoyens de la politique pour comprendre les enjeux des choix collectifs à travers, notamment, la visualisation de données (Bonaccorsi, 2016 :

121). Nous souhaitons pour notre part nous focaliser sur l'apport des données ouvertes comme objets de savoirs et sur la pédagogie à construire avec les élèves autour de ces objets.

L'éducation comme enjeu d'acculturation et horizon d'attente dans la gouvernance des données

Le projet de partage des données s'inscrit dans une logique démocratique et citoyenne considérant les données comme un patrimoine commun. Poser les fondements de la connaissance sur les données apparaît alors à la fois essentiel pour faire des choix de gouvernance en matière de gestion, de normalisation, de médiation et de diffusion, mais implique que des acteurs s'en saisissent et réfléchissent aux moyens permettant de développer l'acculturation du public.

Dans les administrations, le besoin de sensibilisation et de formation des personnels à tous les niveaux commence à se faire sentir pour sortir des fonctionnements en silos (par métiers) et croiser les disciplines ou les compétences. Dans l'éducation, la majorité des acteurs ne semble pas encore se sentir concernée. C'est plutôt par le biais des associations, telles que Fréquence Ecoles ou la FING (Fédération Internet Nouvelles Générations) ou au niveau du CLEMI (Centre de Liaison des Médias d'Information) du côté du Ministère de l'Éducation Nationale, que l'on voit se constituer des groupes d'enseignants passionnés par le numérique qui réalisent des projets avec leurs élèves, soutenus et encouragés par ces organismes. Dans les entreprises, la gouvernance de l'information suppose des choix de gestion, d'évaluation et de contrôle de la qualité des données, dans une perspective de gestion du risque. Elle se met en place dans une logique hiérarchique, même si elle suppose un partage de principes communs dans les équipes. Dans le cas des données publiques ouvertes, les problématiques de gouvernance sont complexes, parce que les objectifs de l'ouverture ne relèvent pas uniquement de logiques économiques, ni de protection, au contraire. Le souci de valorisation et de médiation des données prime sur celui de la protection, puisque l'une des valeurs fondatrices de l'ouverture est la transparence. Du point de vue organisationnel, la gouvernance devrait fonctionner sur le mode participatif et plutôt "bottom-up". Si l'on considère la question de l'éducation, il s'agira alors de se demander dans quelle mesure les acteurs de l'éducation, administration, enseignants, mais aussi apprenants, peuvent participer à un écosystème global et contribuer aux choix de contenus, de formats, de normes, d'indexation, et d'usages des données. On peut alors parler d'une gouvernance participative qui contribuerait également à l'éducation des élèves en les impliquant à la source de la production des données.

Une approche compréhensive de la question du lien entre l'éducation aux données ouvertes et leurs modes de gouvernance

Afin de cerner la problématique dans sa globalité et avec un recul suffisant, nous avons choisi d'interviewer une pluralité d'acteurs impliqués aux niveaux politique, associatif, militant, ainsi que les utilisateurs en contexte éducatif.

Identification des acteurs impliqués

Nous avons cherché des enseignants qui s'intéressent et utilisent les données ouvertes avec leurs élèves à partir des projets soutenus par le CLEMI dans l'Académie de Bordeaux et au niveau national à travers les *Travaux Académiques Mutualisés (TRaAM)*, soutenus par la Direction pour le numérique éducatif, concernant l'Éducation aux Médias et à l'Information. Nous avons également interrogé les formateurs de formateurs à l'ESPE d'Aquitaine pour repérer des acteurs impliqués dans cette thématique. Enfin, nous avons identifié sur les réseaux socio-numériques les comptes les plus actifs sur Twitter principalement, parmi les acteurs de l'éducation. Partant du terrain académique, nous avons interviewé les enseignants ainsi que leurs élèves, puis les cadres de l'éducation (Inspecteurs d'Académie - Inspecteurs Pédagogiques Régionaux, Délégués au Numérique et à l'Education aux Médias et à l'information), jusqu'au niveau national, à la mission ministérielle « Numérique et nouvelles pédagogies » et au Conseil supérieur des programmes. Nous nous sommes également attachées à cerner le point de vue des cadres administratifs dans les collectivités territoriales (commune, métropole et département), ainsi que celui des membres de la FING, proactifs sur la question autant via internet qu'auprès des collectivités.

Choix de méthodes

Pour démarrer cette recherche, nous avons mené une enquête auprès des diffuseurs de données d'un côté, notamment à travers les sites, portails, institutions, collectivités, associations, et des usagers de l'autre (institutionnels et enseignants), afin d'identifier les objectifs pédagogiques qu'ils conçoivent et leurs représentations de ces données pour un usage pédagogique et la construction d'une "data literacy". Cette phase d'enquête qualitative, de type ethnographique, a été menée à partir d'entretiens semi-directifs et d'observations auprès d'une dizaine d'acteurs. Elle sera suivie d'un recueil des pratiques sur la base de nouvelles observations sur le terrain et d'entretiens avec les enseignants et les élèves dans les classes qui mènent des projets autour des données ouvertes. Les objectifs de cette première recherche sont d'identifier :

- les représentations des usages potentiels et réels de ces données,
- les stratégies de sélection et de mise à disposition des données, les choix économiques, politiques, organisationnels,
- les objectifs pédagogiques,
- les stratégies de communication et de médiation autour des données.

Pour cela nous avons isolé plusieurs critères :

- 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.

Enjeux de l'éducation aux données ouvertes : les questions émergentes

Les acteurs que nous avons interrogés sont des producteurs/diffuseurs ou des prescripteurs et utilisateurs de données. Selon leur positionnement, leurs définitions des données ouvertes sont variables, même si la plupart portent un discours proactif, voire militant, et en sont conscients, que ce soit dans les collectivités locales ou dans le monde de l'enseignement. A partir de ces discours, quelques enjeux saillants émergent.

Des enjeux politiques : transparence et accès à l'information

Le discours sur les données ouvertes est marqué par l'idéologie de la transparence liée au principe d'accès à l'information. En France, c'est le droit individuel d'accès à l'information qui est affirmé sur le plan des principes dès 1789 dans l'article 15 de la Déclaration des droits de l'homme et du citoyen, sur le *droit de demander compte à tout Agent public de son administration*. L'article 19 de la Déclaration universelle des droits de l'homme de 1948 affirme que *"tout individu a droit à la liberté d'opinion et d'expression, ce qui implique le droit de ne pas être inquiété pour ses opinions et celui de chercher, de recevoir et de répandre, sans considérations de frontières, les informations et les idées par quelque moyen d'expression que ce soit"*. Cependant, sur le plan des moyens réels d'accès à l'information, la libération des données se développe de manière irréversible mais inégale sur les différents territoires français, dans un cadre institutionnel récent avec la mise en place de la mission Etalab (2011). Cette mission est l'aboutissement d'une évolution de l'agenda politique à partir de la loi du 17 juillet 1978, dite loi CADA sur le droit d'accès aux documents administratifs. En Amérique du Nord, le souci de transparence de la démocratie a conduit à l'adoption dès 1966 du Freedom of information Act puis en 2007 de la loi Open Government Act. L'*open data* est donc intrinsèquement lié à l'*open government*. La directive européenne dite PSI sur la réutilisation des informations du secteur public de 2013 a entraîné une révision de la loi CADA en 2015, posant le principe de la gratuité des réutilisations, augmenté du principe d'ouverture par défaut des données publiques dans la loi pour une République numérique de 2016.

Dans les entretiens que nous avons menés, cette référence à la transparence comme horizon incontournable de la démocratie est sous-jacente mais rarement explicitement nommée. C'est du côté du monde enseignant qu'elle est la plus présente. Un enseignant de Sciences Economiques et Sociales voit aux origines du projet d'ouverture des données *« un mouvement militant qui s'est créé pour dire nous avons aussi un droit de regard sur les données publiques produites par les citoyens et donc il y a eu un mouvement de militantisme fort pour essayer de mettre à disposition du public les données qui sont stockées par les pouvoirs publics »*.

Pour les producteurs de données, l'enjeu politique essentiel des données ouvertes est communicationnel. D'un point de vue très pragmatique, ils soulignent le fait que la décision de mettre en place un portail de données ouvertes ou de consacrer des moyens humains à des projets d'ouverture dépend de considérations et de stratégies de communication politique. La transparence est le fil conducteur des discours, mais n'explique que très partiellement les politiques réelles. Les équilibres politiques au niveau local sous-tendent souvent les choix techniques qui sont faits, destinés à valoriser une collectivité, à mettre en relief des stratégies d'alliance aux niveaux politique local ou national, à participer au développement du territoire.

Enfin, les questions d'organisation des services entre eux, de modernisation et de management apparaissent comme essentielles. Travailler sur l'ouverture des données dans une collectivité exige une réorganisation des compétences "métiers" et une acculturation des professionnels et des services à des questions de communication interne et externe. Les projets d'ouverture appellent donc une remise en question des routines bureaucratiques et des identités professionnelles dans la mise en perspective du travail des services avec le monde extérieur et entre les services. Les choix de jeux de données ne sont pas faits uniquement en fonction de considérations de service au public ou de transparence démocratique mais aussi (surtout) en fonction des modes d'organisation du travail au sein des collectivités. L'ouverture des données est donc aussi un moyen de moderniser les services, d'améliorer leur efficacité et de réorienter les métiers à partir d'une problématique sociale et politique, celle de la transparence, mais toujours autour d'une problématique technique d'information et d'informatique. Il s'agit de sortir d'une logique de silos de savoirs et de métiers, en associant l'entrée dans la technique/informatique et l'entrée dans une politique renouvelée. Les collectivités territoriales en France ont des stratégies variables autour de ce levier de changement que sont les données ouvertes, certaines donnant l'impulsion et s'appuyant plutôt sur les politiques (cabinet, élus), sur les directions des services informatiques (DSI), mais d'autres sur celles des ressources humaines, sur le service de communication, sur les services généraux, ou sur des alliances entre services et élus.

On voit donc bien que les enjeux politiques divergent sensiblement entre les producteurs/diffuseurs et les usagers enseignants dans les représentations, même si les bases idéologiques sont communes sur le plan des principes de référence. Ces divergences s'expliquent par les nécessités pragmatiques et les contraintes de l'action collective dans les organisations ainsi que par le jeu des différents acteurs dans les systèmes que constituent les collectivités publiques. Les données publiques représentent d'un côté un puissant instrument de communication politique. De l'autre, les enseignants sont contraints par les nécessités des programmes et focalisés sur des préoccupations éducatives de circulation de l'information et d'approche critique. Les données représentent pour eux un riche objet de savoirs et d'enseignements.

Des enjeux sociaux : participation à travers les usages innovants vs exploitation commerciale

La logique d'ouverture qui garantit la transparence de l'action politique est complétée par une logique de participation, voire de concertation. Le principe mis en avant est celui de la nécessité sociale d'associer le public à la création de l'information dans son contenu, sa mise à disposition, son utilisation, jusque dans la prise de décision qui nécessite une information précise et complète. Ce principe de participation est lié à celui d'innovation, récurrent dans les discours. L'accent est clairement mis dans le discours sur l'accompagnement des usages innovants et l'amélioration des services rendus aux usagers, notamment dans des domaines sensibles comme l'action sociale, l'inclusion des personnes handicapées ou empêchées, l'environnement. Les dénominations des services ou des missions dans les organigrammes des collectivités territoriales sont explicites : "service accélérateur des usages de la direction innovation et aménagement numérique", pour l'une, "service transition numérique" pour l'autre, par exemple. Dans les définitions données par les acteurs des données ouvertes, on trouve ainsi l'utilisation et la réutilisation comme éléments de base et finalité de la mise à disposition des données. Tous les producteurs de données et la plupart des utilisateurs définissent la donnée ouverte par sa caractéristique d'être réutilisable, avec plus ou moins de précision technique sur les conditions de cette réutilisation. Certains mentionnent le crowdsourcing et les outils de contribution collaborative, donc le développement de services permettant d'associer les usagers à la collecte et à l'usage de données, notamment autour de la cartographie (Open Street Map). Cette dimension participative est évoquée par les enseignants également, et même une élève qui a expérimenté la création d'un réseau participatif de collecte de données à visée d'amélioration de services publics, en constatant les lacunes des données mises à disposition par sa ville.

Cependant, la mise en place de réelles politiques participatives autour de l'ouverture reste rare et située dans quelques territoires seulement (la Bretagne, Paris notamment). Les politiques lient plutôt la mise à disposition des données avec le développement économique et l'innovation technologique. Le "grand public" n'est plus spécifiquement la cible qu'indirectement. Les entreprises via le développement d'applications, éventuellement les journalistes dans le cadre du data journalisme, sont plutôt considérés comme les utilisateurs potentiels les plus crédibles, dans une perspective réaliste. L'utilisation de données ouvertes nécessite en effet des compétences spécifiques qui sont développées dans les organisations. Celles-ci, avec les services internes des collectivités, sont donc plutôt visées, y compris dans les associations militantes comme la FING, qui oriente nettement son action vers le développement de compétences "métiers" liées aux usages des données ouvertes. Un nouveau métier de *data scientist* émerge ainsi. Dans cette perspective, les formes de mise à disposition des données visent des usages experts et destinés à dégager des profits économiques ou sociaux. Même dans le cadre des communautés du libre, une chargée de mission souligne la complexité des logiciels libres comme Open Street Map et la nécessité d'une expertise pour les utiliser, voire le fait que cette complexité frôlant l'esotérisme est volontairement entretenue dans le but d'éloigner des contributeurs novices incompetents ou inopportuns et de conserver une identité au groupe expert des usagers.

Sur le thème de la participation, les discours des producteurs et des enseignants divergent profondément. Pour les enseignants de sciences économiques notamment, les usages commerciaux des données ouvertes représentent un danger et un risque. Un danger, parce qu'il s'agit d'une forme de prédation par rapport à l'action publique de mise à disposition de données, un risque, parce que les usages commerciaux peuvent porter atteinte à des droits fondamentaux comme le respect de la vie privée quand des données anonymes peuvent être croisées et permettre soit d'identifier les personnes en jeu, soit de créer des outils de manipulation, dans une *"dérive qui passe du maillage fin de la population au ciblage du consommateur puis du citoyen"*. Pour les acteurs des collectivités locales, la possible exploitation industrielle et commerciale des données, notamment par les entreprises locales, est avant tout perçue comme un levier pour la croissance et l'innovation. Ainsi, certains acteurs militants des collectivités sont prêts à ouvrir toutes les données, à faire confiance en effectuant des actions d'animation visant leur utilisation large, sans même s'autoriser à anticiper ni à présupposer les réutilisations qui pourraient en être faites.

Des enjeux normatifs et documentaires : mise à disposition et organisation des données

Nous avons vu que la réutilisation des données est, pour la plupart des acteurs, la caractéristique fondamentale de la donnée ouverte. Trois éléments sont essentiels pour la réutilisation : les licences, les formats, les métadonnées. Le premier détermine les conditions juridiques de la réutilisation, et résulte de choix politiques ; le second les conditions techniques et résulte des conditions de production des données par les services ainsi que des systèmes informatiques ; le troisième les conditions documentaires et résulte également des services qui intègrent l'importance de l'indexation notamment. La question des licences est déterminante pour comprendre les options fondamentales choisies par les acteurs quant au rôle assigné à l'ouverture des données dans une éventuelle redistribution du pouvoir aux citoyens sous l'action des collectivités publiques (Peugeot, 2012). Les licences ouvertes sont nombreuses actuellement et varient essentiellement selon l'obligation faite à l'utilisateur soit de reverser les données créées sous les mêmes conditions d'ouverture (obligation de réciprocité comme dans la licence ODbL de l'Open Knowledge Foundation) soit de seulement citer la source, encourageant ainsi une libre réappropriation des données (comme dans la licence ouverte OL). L'absence de contrainte de réciprocité, dans une logique libertaire et libérale, est vue comme un accélérateur des usages qui favorise la création et le développement de services et d'applications par les entreprises. La condition de réciprocité est une position de principe inspirée de la communauté du libre dans laquelle l'ouverture et la liberté sont la règle et des obligations à l'usage.

Ces points juridiques sont essentiels et n'échappent pas aux enseignants. Pour les acteurs de l'enseignement des Sciences Economiques et Sociales, la domination du marché des données numériques par les entreprises et plus particulièrement par les grandes entreprises du numérique (Google, Amazon, Facebook, Apple, Microsoft) à travers l'usage de puissants algorithmes,

présente un enjeu majeur. Une méfiance profonde s'exprime donc chez eux, notamment par rapport aux usages possibles et asymétriques des données ouvertes. Dans cette perspective, les données ouvertes sont associées à la thématique du *big data* et à des enjeux citoyens. Cette insistance sur les risques d'asymétrie soulignés dans la littérature scientifique (Chartron, Broudoux, 2015) est propre aux enseignants de sciences économiques et sociales. Pour les enseignants plus précisément chargés de l'EMI, c'est la "*culture du bien commun*" associée à l'action de "*l'intelligence collective*" qui est valorisée.

Le point technique du format de mise à disposition est souligné par tous les acteurs, producteurs et usagers. Pour les producteurs, la question du format est gérée dans une négociation entre les différents services et la DSI. Cette négociation peut être longue, parce qu'elle dépend des historiques des différents services et des usages mis en place. L'ouverture des données, qui nécessite identification, extraction, et "brutification" (Denis, Goéta, 2013), représente souvent un accélérateur des changements dans ce domaine encore, puisque la nécessité de réutilisabilité et d'interopérabilité est prise en compte par les services. Les enseignants, dès qu'ils ont mis en place des projets, connaissent les différents formats, leurs contraintes et leurs avantages. Ils évoquent tous le format .csv et savent que la création d'information nécessite des compétences techniques liées au formatage des jeux de données.

Enfin, l'utilisabilité des données ouvertes dépend en grande partie de la qualité des métadonnées, qui permet la contextualisation des jeux de données, comme le souligne un responsable de collectivité territoriale que sa formation d'archiviste a sensibilisé à l'importance des informations contextuelles. Charlotte Maday (2015 : 161) souligne ainsi la proximité de la gestion des données ouvertes avec le *record management* et la production de métadonnées de description, de gestion et techniques. Dans les administrations engagées dans la mise en place de plateformes de données ouvertes, bien que tous s'orientent souvent vers le même système de gestion de données de référence en la matière, qui propose les mêmes éléments de métadonnées (à savoir la plateforme britannique CKAN), les schémas de classifications et d'organisation des métadonnées sont hétérogènes sur les différents territoires (Ben Abdallah, 2016 : 52). Les choix des métadonnées à renseigner ne sont pas réfléchis collectivement alors qu'ils ont pourtant une importance capitale et dépendent de la stratégie visée. Celle-ci peut être plutôt d'ordre documentaire avec des métadonnées descriptives permettant aux usagers de trouver plus facilement les jeux qu'ils cherchent, ou plutôt d'ordre archivistique avec des métadonnées visant l'intégrité et la pérennité des données dans un but d'archivage (ibid. : 44). Au coeur du projet de gouvernance des données, cette question de l'indexation et de la contextualisation des ressources ne semble intéresser que certains professionnels, mais pas tous les responsables, notamment des DSI. Au département de la Gironde par exemple, le chef de projet documentaire s'est occupé seul de cette question, qui n'apparaît pas comme une priorité ou un enjeu majeur de la stratégie politique.

Les enseignants ou élèves utilisateurs interrogés nous ont dit ne pas prendre part à la création de métadonnées ni à l'indexation des données disponibles en ligne, ne sachant d'ailleurs pas toujours si cela est possible sur les plateformes utilisées. Cette question ne représente pas non plus un enjeu pour eux, sans doute car elle est trop technique. En revanche, leur contribution par la création de nouvelles données, pour les rendre exploitables par d'autres, ou pour construire du sens et communiquer une information, apparaît comme un point d'entrée susceptible de favoriser la concertation entre les acteurs et un mode de gouvernance par le bas. Dans ce cas, la question des métadonnées qui nous paraît être un enjeu important pour favoriser la réutilisation des données, pourrait être abordée au moment de la formation.

On voit ainsi émerger chez les enseignants des discours utopiques ou dystopiques en décalage par rapport aux producteurs de données qui s'attachent plus précisément à des dimensions normatives, techniques ou documentaires.

Des enjeux éducatifs : médiation, animation, formation et lien entre diffuseurs et utilisateurs de données

Les agents des collectivités parlent essentiellement d'animation. Si la formation des élèves leur paraît tout à fait intéressante et nécessaire, ils ne cherchent que rarement directement des formes de concertation avec l'éducation. Les acteurs impliqués dans l'animation de réseaux d'utilisateurs des données soulignent l'importance de la formation du public, ainsi que des agents métiers dans les collectivités, comme condition indispensable à l'ouverture des données. Certains acteurs des collectivités territoriales sont très impliqués dans des actions de formation, autour de la cartographie par exemple, plutôt dans une perspective d'animation de quartiers. Les associations militantes ont mis en place des formes de collaboration avec le monde éducatif, avec des actions de formation, et prennent en compte la nécessité de diffuser les connaissances et les compétences sur les données ouvertes dans l'enseignement.

Les enseignants ne sont pas des militants engagés dans les politiques d'ouverture des données. Ils adoptent cependant tous des démarches actives de veille individuelle, en l'absence d'un réseau structuré parmi les enseignants sur ce thème et d'impulsion au niveau institutionnel. Ils sont autodidactes et tous fortement impliqués dans l'éducation aux médias et à l'information, en lien avec leur discipline. L'offre de données leur apparaît encore trop limitée pour être créatrice de connaissances dans le cadre des programmes. Un enseignant souligne que l'intérêt de l'usage des données ouvertes est son ancrage dans la "vraie vie", avec la création d'informations à partir de données réelles. L'enseignant de technologie insiste sur la démarche de formulation d'hypothèses, la recherche de données pour des statistiques, le traitement avec des outils comme le tableur, la mise en forme par la visualisation dans une perspective de communication et de publication. Dans une perspective pluridisciplinaire, comme les projets des travaux personnels encadrés en classe de première, les données sont exploitées dans un cadre large de questions liées aux programmes et de l'éducation aux médias et plutôt dans une démarche de projet. Cependant, la thématique des données ouvertes reste peu

attractive pour l'ensemble des enseignants qui ignorent très souvent son intérêt ou redoutent sa complexité dans l'usage de bases de données.

L'impulsion des projets de formation des enseignants vient clairement de la hiérarchie et notamment de la Direction du numérique pour l'éducation, mais à partir d'expériences d'enseignants sur le terrain qui se sont emparés de la thématique des données ouvertes. Cette impulsion offre l'occasion d'inciter les enseignants à travailler en interdisciplinarité autour de projets, et à partir d'une entrée liée au numérique (par exemple la nouvelle option Informatique et Culture Numérique). Pour les enseignants en sciences économiques sociales, le travail autour des données concerne trois grandes thématiques centrées sur le "société numérique" : l'origine des données (avec une référence à Dominique Cardon), le traitement des données (et notamment les algorithmes), les utilisations et la diffusion des données (avec l'infographie), en lien avec la protection des individus et l'approche critique de l'information, par exemple autour des questions de complot. Un certain flou plane sur la spécificité des données ouvertes, et, comme dans la presse (Goéta, 2015), on peut noter une confusion entre *open data* et *big data* : il s'agit toujours de travailler sur les données. La thématique des données ouvertes en éducation permet donc de concilier une approche technique et informatique avec une approche culturelle et critique des questions informationnelles.

Conclusion : Vers un changement de paradigme ?

Les discours sur les données ouvertes révèlent une multiplicité d'enjeux et des contradictions qui produisent progressivement des effets. Si le consensus autour de l'ouverture des données est général, le principe de transparence inattaquable, les divergences se révèlent très vite dans la réalité de la mise en oeuvre, qui pointe l'importance de choix stratégiques et politiques (entre libéralisme, libertarisme et étatismes par exemple) et de contraintes organisationnelles majeures. L'usabilité des données mises à disposition apparaît inséparable de la mise en place de moyens autour de la capacité du public à s'en saisir, la *data literacy*, si on veut éviter que ces données ne profitent qu'aux entreprises qui ont les moyens de les traiter. On voit donc se mettre en place un déplacement des compétences et des connaissances à partir de la prise de conscience de la centralité des données, de leur appropriation, des usages possibles. Si l'on revient à la question de la gouvernance, il semble que les données ouvertes soient devenues assez rapidement un enjeu collectif autour duquel des communautés plurielles se sont mobilisées, pour des raisons diverses. L'objectif de transparence démocratique et de formation du citoyen n'est pas nécessairement primordial dans la prise de décision politique, mais les projets d'ouverture des données permettent de sensibiliser les administrations et les acteurs de l'éducation sur la question englobante des données. On ne peut cependant pas parler aujourd'hui de gouvernance participative, parce que les enjeux restent encore flous et discordants entre les acteurs. L'éducation aux données, qui reste à construire, revêt finalement trois dimensions essentielles :

- une culture numérique pour trouver et manipuler les données, les algorithmes, les outils de visualisation ;
- une culture informationnelle pour comprendre l'importance des données et savoir passer de la donnée à l'information, de l'information à la connaissance ;
- une culture politique centrée sur le droit à l'information, la transparence, la maîtrise des données en général et des données personnelles en particulier, sur la participation, et sur la gouvernance par les citoyens.

Bibliographie

Ackoff R.L., 1989. « From Data to Wisdom », *Journal of Applied Systems Analysis*, 16, pp.3-9.

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.

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., 2013, *Médiactivistes*, 2e éd., Paris, Presses de Sciences Po (P.F.N.S.P.).

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, pp. 67-83.

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

Denis J., Goëta S., 2013, « La fabrique des données brutes. Le travail en coulisses de l'open data », in : Journée d'études SACRED Penser l'écosystème des données. Les enjeux scientifiques et politiques des données numériques, Paris - 13 février, [En ligne : <https://halshs.archives-ouvertes.fr/halshs-00990771/document>]

Doueihy M., 2011, *La grande conversion numérique*, Paris, Seuil.

El Hachani M., 2015, « Open data, collectivités et usagers : une dynamique en question », in : Paquienséguy F., dir., *Open data. Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, Paris, Editions des archives contemporaines, pp. 1-23

Goëta S., 2015, « Un air de famille : les trajectoires parallèles de l'open data et du big data », *Informations sociales*, 5, n° 191, pp. 26-34.

Guyot B., 2013, « La gouvernance de l'information, point de rencontre complexe entre stratégie et transversalité », in Guyot B., Fuzeau P., Cottin M., Rietsch J.-M.

dir., Repères conceptuels, *Documentaliste-Sciences de l'Information*, 1, Vol. 50, pp. 26-29.

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

Kembellec Gérald, Prime-Claverie Camille, 2016, « Web de données et création de valeurs : le champ des possibles », *I2D – Information, données & documents* 2, Vol. 53, pp. 28-29.

[En ligne : : www.cairn.info/revue-i2d-information-donnees-et-documents-2016-2-page-28.htm.], consulté le 18/08/2016.

Labelle S., Le Corf J-B., 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 Boisnier S. et Pailliar I. (coord.), *Les Enjeux de l'Information et de la Communication*, n° 13, Vol. 2, pp. 59-71.

Maday C., 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*, 11, Vol. 2, pp. 149-166.

Maurel D., Chebbi A., 2012, « La perception de la confiance informationnelle », *Communication et organisation*, 42 [En ligne : <http://communicationorganisation.revues.org/3864>], consulté le 14/07/2016.

Meijer A., Curtin D., Hillebrandt M., 2012, « La gouvernance ouverte : relier visibilité et moyens d'expression », *Revue Internationale des Sciences Administratives*, 1, Vol. 78, pp. 13-32.

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.», [En ligne http://archivesic.ccsd.cnrs.fr/sic_00759618].

Open Data France, 2016, *Guide Open Data pour les communes / Glossaire de la donnée publique*, [En ligne : http://www.opendatafrance.net/wp-content/uploads/2016/06/guideOD_communes_glossaire_juin2016_Web.pdf].

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*, 1, Vol. 73, pp. 163-179.

Peugeot V., 2011, « L'ouverture des données publiques : convergence ou malentendu politique ? » , in Stiegler B. dir., *Confiance, croyance, crédit dans les mondes industriels*, Paris, FYP, pp. 79-120.

Proulx S., 2015, « Usages participatifs des technologies et désir d'émancipation : une articulation fragile et paradoxale », *Communiquer*, 13, pp. 67-77.

Schafer V., Le Crosnier H., 2011, *La neutralité de l'Internet: un enjeu de communication*, Paris, CNRS Editions.