



HAL
open science

La réutilisation et les réutilisateurs des données ouvertes en France: une approche centrée sur les usagers

Valentyna Dymytrova, Françoise Paquienséguy

► To cite this version:

Valentyna Dymytrova, Françoise Paquienséguy. La réutilisation et les réutilisateurs des données ouvertes en France: une approche centrée sur les usagers. *Revue Internationale des Gouvernements Ouverts*, 2017. hal-01718057

HAL Id: hal-01718057

<https://hal.science/hal-01718057>

Submitted on 27 Feb 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.

Comment citer cet article

DYMYTROVA, Valentyna; PAQUIENSÉGUY, Françoise. La réutilisation et les réutilisateurs des données ouvertes en France: une approche centrée sur les usagers. **Revue Internationale des Gouvernements Ouverts**, [S.l.], v. 5, p. 117-132, août 2017. ISSN 2553-6869. Disponible à l'adresse : ><http://ojs.imodev.org/index.php/RIGO/article/view/204>>.

LA RÉUTILISATION ET LES RÉUTILISATEURS DES DONNÉES OUVERTES EN FRANCE : UNE APPROCHE CENTRÉE SUR LES USAGERS

Valentyna Dymytrova, Post-doctorante en Sciences de l'Information et de la Communication, EA 4147 ELICO, Université Lyon 1

Françoise Paquienséguy, Professeur en Sciences de l'Information et de la Communication, EA 4147 ELICO, Sciences Po Lyon

Bien qu'elle soit un des éléments clefs des discours d'accompagnement de l'ouverture des données, la notion de réutilisation reste mal définie aussi bien par les acteurs publics que par les entreprises du numérique ou les chercheurs, chacun investissant cette notion d'objectifs et d'attentes différents. Ces discours sur les données ouvertes ou sur l'ouverture des données font de la réutilisation à la fois un outil témoignant d'une plus grande transparence de l'administration et de sa modernisation, et un facteur de développement des villes intelligentes, assurant une croissance économique et une innovation technologique sensées profiter aux citoyens¹. Toutefois, malgré un nombre important de jeux de données publiques rendues disponibles ces dernières années *via* des portails métropolitains, le nombre d'applications reste assez limité², tout autant que celui d'utilisateurs car « le nombre de services qui parviennent à la viabilité et la pérennité l'est encore plus »³. D'ailleurs, il existe peu d'études empiriques sur la réutilisation des données ouvertes sur les portails dédiés⁴ et l'intérêt porte plus souvent sur les standards techniques de l'interopérabilité que sur les horizons d'attentes des réutilisateurs. À la différence de ces modèles dominants d'analyse, centralisés et « top-down », notre recherche est guidée par une approche centrée sur l'utilisateur⁵, ses difficultés, besoins et attentes liés à la réutilisation des données ouvertes. Autrement dit : qu'en est-il des réutilisateurs, usagers intermédiaires des données ouvertes ?

En mobilisant les résultats d'une enquête de terrain menée dans le cadre de l' ANR-14-CE24-0029⁶, cet article vise à identifier les profils des réutilisateurs, leurs stratégies et pratiques afin de mieux cerner les leviers et les difficultés de la réutilisation des données ouvertes car ces enjeux

¹ S. CHIGNARD, *Open data : Comprendre l'ouverture des données publiques*, Paris, FYP, 2012 ; F. PAQUIENSÉGUY (dir.), *Open data : Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, Paris, Editions des Archives contemporaines, 2016.

² <http://cordis.europa.eu/fp7/ict/content-knowledge/docs/open-data-reuse-incubator4-3-a.pdf>.

³ S. TURKI, M. FOULONNEAU, « Valorisation des données ouvertes : acteurs, enjeux et modèles d'affaires », in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. BROUDOUX, G. CHARTRON (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 113.

⁴ R. KITCHIN, *The Data Revolution: Big data, Open data, data infrastructures and their consequences*, London, Sage, 2014.

⁵ D. A., NORMAN, S.W. DRAPER (Eds.). *User-centered system design: New perspectives on human-computer interaction*, Hillsdale, NJ, Lawrence Erlbaum, 1986.

⁶ « Fostering Uses and Usages of Open Sensor Data in Smart Cities »: <http://opensensingcity.emse.fr/>.

sont construits *a posteriori*, une fois l'ouverture effective. Celle-ci sera ici comprise comme une stratégie d'offre, de mise à disposition par les collectivités territoriales et tout particulièrement les métropoles françaises instituées au 1^{er} janvier 2015, *via* leurs portails opendata dont les usages n'ont été ni conçus ou anticipés mais dictés par l'injonction publique⁷.

Pour cerner la question, nous préciserons d'abord les significations des termes « réutilisation » et « réutilisateurs » dans les textes législatifs et les documents administratifs et au sein des portails métropolitains opendata (PMOD). Ensuite, nous analyserons les représentations et stratégies associées à la réutilisation des données ouvertes dans les pratiques professionnelles des réutilisateurs interviewés dans le cadre de notre enquête (n=22). Enfin, nous dégagerons les décalages et les tensions entre les différents discours et les pratiques professionnelles analysées.

§ 1 – QU'EST-CE QUE LA RÉUTILISATION DES DONNÉES OUVERTES ?

Surtout caractéristique d'un contexte technique, le terme « réutilisation » figure dans la plupart des définitions d'ouverture des données, à l'instar de celle proposée par la fondation Open Knowledge International : « une donnée ouverte est une donnée qui peut être librement utilisée, réutilisée et redistribuée par quiconque - sujette seulement, au plus, à une exigence d'attribution et de partage à l'identique »⁸. La réutilisation est du registre de l'idéologie et doit être reliée à deux des trois grandes utopies contemporaines du web⁹ : l'ouverture et la participation. Ces idéaux laisseraient penser que le lien se crée de façon automatique ou spontanée entre offre et usages, entre données ouvertes et réutilisations, alors que, nous le savons, mais nous le verrons aussi, ces processus doivent se construire et les pratiques professionnelles coloniser ces nouveaux espaces que sont les PMOD (portails métropolitains opendata). Ces deux maîtres mots d'ouverture et de participation se trouvent pris dans un contexte contraignant à plus d'un titre. Ainsi les cadres, au sens de Flichy¹⁰, technique, juridique et d'usages sont-ils aujourd'hui terriblement restrictifs, ils peinent à converger pour construire ensemble le cadre de référence qui témoignerait de la mise en œuvre d'un véritable écosystème de la réutilisation. On remarque d'ailleurs très vite dans les propos des professionnels interrogés que la réutilisation constitue l'une des caractéristiques idéales des données ouvertes au même titre que l'accessibilité ou l'absence de restrictions techniques¹¹. Ne serait-elle qu'un horizon d'attente ? Utopique ? Bien sûr, les données ouvertes sont censées être réutilisables techniquement (disponibilité dans un format pratique et modifiable), juridiquement (sous une licence autorisant leur réutilisation et redistribution incluant le croisement avec d'autres ensembles de données), économiquement (possibilité de redevances) et socialement (tout le monde doit être en mesure de réutiliser et redistribuer les données ouvertes), mais la réalité est-elle conforme à l'injonction de réutilisation, autrement dit d'appropriation, d'acculturation de la donnée et de ses jeux ?

A) La réutilisation, caractéristique utopique de l'opendata ?

Le projet public d'ouverture des données institue leur réutilisation qui représente l'objectif et la raison d'être de la mise à disposition, comme le souligne la Charte du G8 pour l'Ouverture des Données Publiques,¹² qui voit dans la réutilisation, la possibilité d'« acquérir des connaissances,

⁷ L'enquête conduite au nom d'Ellico EA 4147, par Valérie Larroche-Boutet, Martine Vila, Marie-France Peyrelong, Valentyna Dymytriva et Françoise Paquienséguy qui en a aussi assuré la direction scientifique. Le panel d'interviewés est constitué de réutilisateurs inscrits et déclarés comme tels sur des portails métropolitains. Au total, 22 entretiens ont été menés de septembre 2015 à juin 2016.

⁸ OpenDataHandbook par Open Knowledge International: <http://opendatahandbook.org/guide/fr/what-is-open-data/>.

⁹ F. IBEKWÉ-SANJUAN, F. PAQUIENSÉGUY, « Open, Big, Collaboration : trois utopies de l'innovation au 21^e siècle », in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. Broudoux, G. Chartron (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 15.

¹⁰ P. FLICHY, *L'innovation technique. Récents développements en sciences sociales. Vers une nouvelle théorie de l'innovation*, Paris, La Découverte, 1995.

¹¹ R. KITCHIN, *The Data Revolution: Big data, Open data, data infrastructures and their consequences*, London, Sage, 2014, p.50.

¹² Signée le 18 juin 2013 lors du Sommet de Lough Erne, en Irlande du Nord et disponible ici : <http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/charte-g8-ouverture-donnees-publiques-fr.pdf>.

générer des idées et produire des services afin de créer un monde meilleur pour tous ». De même, bien des discours d'accompagnement de l'ouverture de données insistent sur la « réutilisation la plus large possible » et « par le plus grand nombre »¹³. Ainsi, le Vade-mecum sur l'ouverture et le partage de données publiques élaboré par Etalab en 2013 précise que « la réutilisation des données publiques peut susciter le développement de nouveaux services comme les applications mobiles, des sites Internet, des visualisations données ou « datavisualisation » notamment par la presse, etc. » Finalement, l'Administration se met en ordre de marche pour rendre certains jeux de données disponibles (et il y aurait fort à dire sur la désignation de ces derniers) et le reste suivra... Cependant, si les usages sont peu anticipés et de façon extrêmement confuse comme on vient de le dire, les usagers sont définis au regard de la collectivité territoriale et de ses interlocuteurs traditionnels. « Les données publiques peuvent être aussi réutilisées par les chercheurs, les enseignants, les étudiants, les responsables associatifs, les citoyens, pour construire de nouveaux points de vue sur la société ou sur l'action publique »¹⁴. Cependant, plusieurs documents officiels¹⁵ opèrent une distinction entre, d'une part, développeurs et sociétés concevant des applications et, de l'autre, citoyens et urbains, représentés comme de simples usagers-consommateurs de ces dernières¹⁶. Le premier trait de l'intermédiation assurée par les réutilisateurs apparaît ici avec les data-journalistes et les producteurs d'applications. D'ailleurs, certains PMOD proposent aujourd'hui dès leur page d'accueil un accès spécifique aux « geek », « développeur » ou autres « réutilisateurs », il témoigne de leur accompagnement vers une réutilisation. D'autres, la prolongent encore et proposent d'héberger les applications développées à partir de leurs propres données ouvertes, c'est le cas de PMOD de Séoul, de Montréal et d'Amsterdam. Cette réutilisation des données ouvertes par le secteur privé se perçoit comme génératrice d'innovation et de valeur économique et sociale, comme le souligne la Charte du G8 pour l'Ouverture des Données Publiques : « Des données publiques librement accessibles et gratuitement réutilisables peuvent être à la source de services et des produits innovants susceptibles d'aider les individus à faire face plus facilement à la vie moderne. Utilisées de la sorte, ces données amènent le secteur privé à innover, contribuant à la création de nouveaux marchés, de nouvelles entreprises et de nouveaux emplois »¹⁷.

Autrement dit, la mise en œuvre de la réutilisation n'est pas souvent soutenue par l'action publique d'ouverture, dans un premier temps du moins. La tâche est ardue et la réforme territoriale la complique par sa redistribution des compétences d'une part et son lot de réorganisations¹⁸, de l'autre. Comme en phase d'innovation, tout reste à la fois à inventer et à produire en termes de standards techniques certes, mais surtout de compétences, de gouvernance, de métiers et surtout d'écosystème informationnel. Et c'est bien là la dernière visée utopique portée par l'ouverture des données : générer une smart-city, le double numérique et computationnel d'un espace physique, qu'il faut apprendre à habiter¹⁹. Poser un cadre juridique peut être un des premiers éléments pour l'entreprendre.

¹³ Vade-mecum sur l'ouverture et le partage de données publiques. Etalab. Septembre 2013. En ligne : <http://www.enssib.fr/bibliotheque-numerique/documents/61618-vade-mecum-sur-l-ouverture-et-le-partage-des-donnees-publiques.pdf>.

¹⁴ Vade-mecum sur l'ouverture et le partage de données publiques. Etalab. Septembre 2013. En ligne : <http://www.enssib.fr/bibliotheque-numerique/documents/61618-vade-mecum-sur-l-ouverture-et-le-partage-des-donnees-publiques.pdf>.

¹⁵ F. PAQUIENSÉGUY, « Smart city & open data : à qui profitent les données ouvertes ? », *CIST2016 - En quête de territoire(s) ?*, Mars 2016, Grenoble, France. Proceedings du 3e colloque international du CIST, p.351. <https://hal.archives-ouvertes.fr/hal-01353651/>.

¹⁶https://www.europeandataportal.eu/sites/default/files/d2.1.2_training_module_1.5_promoting_the_reuse_of_open_data_through_odip_en_cdp.pdf

¹⁷ <http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/charte-g8-ouverture-donnees-publiques-fr.pdf>

¹⁸ Par exemple, Vade-mecum sur l'ouverture et le partage de données publiques, Etalab, septembre 2013 ; Rapport Trojette sur l'ouverture des données publique, 2013 ; Rapport sur l'ouverture des données de transport présidé par Francis Jutand, mars 2015.

¹⁹ « Smart City – « Le défi : créer de l'intelligence collective ». Entretien avec D. Boullier : <http://www.lemag-numerique.com/2016/01/smart-city-le-defi-creer-de-lintelligence-collective-8337>.

B) Le cadre juridique de la réutilisation

En jurisprudence, la réutilisation renvoie à l'utilisation « à d'autres fins que celles de la mission de service public en vue de laquelle les documents ont été élaborés ou sont détenus »²⁰. Le terme « réutilisation » s'est imposé dans l'écosystème français avec la directive européenne dite « ISP », sur la réutilisation des informations détenues par le secteur public²¹. Elle définit, entre autres, les domaines dans lesquels la réutilisation des informations publiques est autorisée et ses conditions, notamment, l'obligation de transparence sur le calcul des redevances. La présence d'une redevance doit se justifier par des coûts liés à la recherche et au traitement des données engagées par l'administration.

Une autre directive européenne, dite « INSPIRE »²² a donné à son tour une forte impulsion à l'ouverture des données environnementales géographiques et à l'harmonisation des systèmes d'information géographiques au niveau européen en les rendant obligatoires et en définissant les conditions de mise en œuvre²³. En parallèle, la mise à disposition des données publiques et leur réutilisation ont été fortement encadrées par plusieurs dispositifs législatifs nationaux récents (loi Macron, 2015, loi NOTRe, 2015, loi Valter, 2015 et loi Lemaire, 2016). Ceux-ci témoignent d'une formalisation du cadre juridique qui laisse derrière lui une logique de la demande, en référence au droit démocratique de l'accès aux documents administratifs²⁴ au profit d'une logique d'offre que nous avons déjà soulignée et qui n'est en rien dans les traditions et usages administratifs. Elle se fonde sur le droit économique et social de réutilisation des données publiques posée par la directive ISP, transposée en droit français par l'ordonnance du 6 juin 2005²⁵ et le décret du 30 décembre 2005²⁶. Ici, la notion de réutilisation renvoie à l'exploitation des données publiques par des tiers à d'autres fins que celles pour lesquelles elles ont été produites ou reçues, y compris à des fins commerciales, avec une obligation de non-dénaturation et de référencement des informations. En continuité, le décret n°2011-577 du 26 mai 2011 et la circulaire du Premier ministre, F. Fillon, ont fixé le principe d'une réutilisation libre, facile et gratuite des informations publiques pour tous les citoyens. Ces prises nettes de position, en vue de la construction progressive d'un cadre de référence sous la forme d'un écosystème de l'OD, ont pris forme dans la plateforme nationale d'hébergement et de partage de données publiques, ouvertes et réutilisables, produites par l'Etat et les collectivités : data.gouv.fr. Plus récemment, la loi n° 2016-1321 du 7 octobre 2016, dite loi Lemaire affirme le principe de l'ouverture des données publiques par défaut. Elle oblige toutes les administrations à communiquer leurs documents au format électronique, à les mettre à disposition du citoyen dans un standard ouvert, aisément réutilisable et exploitable par un traitement automatisé.

Ainsi, le cadre législatif français promeut-il une réutilisation des informations publiques²⁷, définie en fonction de plusieurs grands principes tels que la liberté et la gratuité des données, laquelle réutilisation s'étend à toutes les données du secteur public ne faisant pas l'objet d'une dérogation et aux données produites par des acteurs privés chargés d'une mission de délégation du service

²⁰ <http://www.cada.fr/qu-est-ce-qu-une-reutilisation,6116.html>.

²¹ Directive 2003/98/EC du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public.

²² Directive 2007/2/CE du Parlement européen et du Conseil établissant une infrastructure d'information géographique dans la Communauté européenne.

²³ D'ailleurs nous avons remarqué dans notre analyse des PMOD que certains avaient structuré leurs rubriques thématiques en reprenant *stricto sensu* celles de la directive, c'est le cas de Lyon par exemple.

²⁴ Loi 78-753 du 17 juillet 1978, dite loi CADA.

²⁵ Ordonnance n° 2005-650 du 6 juin 2005 relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques.

²⁶ Décret n°2005-1755 du 30 décembre 2005 relatif à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques, pris pour l'application de la loi n° 78-753 du 17 juillet 1978.

²⁷ À noter une terminologie variée utilisée selon les sources : données publiques, informations du secteur public ou informations publiques. Comme le souligne le rapport Trojette, « Au regard des normes en vigueur, les expressions « donnée publique », « information publique » et « information du secteur public » semblent devoir être considérées comme synonymes (Rapport Trojette, juillet 2013, p.16).

public à caractère industriel ou commercial. Plusieurs entreprises de ce type commencent d'ailleurs à ouvrir certains jeux de données, par exemple la SNCF, la RATP, Véolia, Decaux ou Transdev. Cette promotion se matérialise particulièrement *via* le lieu de la réutilisation : le portail métropolitain opendata qui donne accès aux jeux de données et, parfois, héberge les applications qui en sont issues.

§ 2 – LE PORTAIL MÉTROPOLITAIN DE L'OPENDATA, PROMOTEUR DE LA RÉUTILISATION

Bien que nous ayons à ce jour analysés 24 PMOD français, européens ou internationaux dans le cadre de l'ANR Opensing city²⁸, le focus premier sera fait ici sur celui de la Métropole de Lyon. En effet, les choix stratégiques qui ont présidé à la genèse et au déploiement de ce portail ont immédiatement fait une place particulière aux réutilisateurs, intégrés dès le back office d'une part et dans la gouvernance de l'autre puisqu'en avril 2015, afin de garantir une bonne gouvernance du PMOD et de faire converger politique et technique, un poste de *Chief Data Officer* a été créé par la Métropole avec comme mission centrale de développer et d'animer le réseau des réutilisateurs, en articulation étroite d'une part avec le réseau professionnel des géomaticiens du territoire et de l'autre avec le LivingLab métropolitain, Le Tuba²⁹.

Ouvert le 26 mai 2013, le portail des données ouvertes de la communauté urbaine de Lyon, *Smart Data Lyon* est devenu *Data Grand Lyon* dans le contexte de la réforme territoriale amenant à la création de la Métropole de Lyon le 1 janvier 2015³⁰. Le PMOD lyonnais est basé sur une plateforme indépendante, développée par les informaticiens de la Métropole elle-même³¹. Elle assure une industrialisation des process d'accès aux données provenant de différents producteurs qui repose sur un modèle économique particulier du courtage informationnel. Mais l'élément à retenir ici est sans doute moins le modèle lui-même que la tentative d'en développer un et de le faire fonctionner afin de stabiliser l'ouverture et la réutilisation dont les coûts de développement, de déploiement et de fonctionnement, supportés par les Métropoles, sont loin d'être négligeables et temporaires. Autrement dit, tenter une structuration en termes de gouvernance et de modèle économique³² témoigne d'une volonté (même inachevée et maladroite) de penser un cadre d'usages stable.

A) Penser la réutilisation

La tâche n'est pas simple malgré les dispositifs législatifs autour de la donnée qui tentent de lever les freins juridiques à la réutilisation et l'innovation des services urbains (obligation de diffusion des données, licences ouvertes, principe de gratuité). Les politiques publiques de la donnée menées par la mission gouvernementale Etalab suivent elles-mêmes les préconisations de l'Union européenne visant à standardiser les politiques en matière d'opendata des pays membres et contribuent à une normalisation et une harmonisation des pratiques. Ce cadre juridique que nous avons déjà évoqué s'avère très prégnant sur l'action locale. « Le développement territorial, qui

²⁸ Analyse disponible dans le livrable n°2 : <https://hal.archives-ouvertes.fr/hal-01449348>.

²⁹ www.tuba-lyon.com.

³⁰ Loi 2014-58 pour la modernisation des territoires et la réaffirmation des Métropoles (loi dite MAPAM, 2014). Ainsi créée, la Métropole de Lyon devient une collectivité territoriale à statut particulier au même titre que les communes, les départements et les régions. Constituée de 59 communes, la Métropole exerce les compétences du Grand Lyon, celles du département du Rhône et a la possibilité d'en acquérir d'autres, par délégation, de la Région ou de l'Etat. Elle a reçu notamment les compétences en matière d'aménagement de l'espace métropolitain (entretien et gestion des routes, transports, voiries).

³¹ Conçu par le service Innovation Numérique & des Systèmes d'Information de la métropole de Lyon, la plateforme technique est entièrement basée sur des logiciels libres : Linux, Apache/Tomcat, PostgreSQL/PostGIS, MapServer 6. Le catalogue des données est basé sur l'outil Géosource.

³² F. PAQUIENSÉGUY, « Les portails Open Data au prisme du courtage informationnel : qu'est-ce qui se joue pour les Métropoles ? », in *Open data : Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, F. PAQUIENSÉGUY (dir.), Paris, Editions des Archives contemporaines, 2016.

repose beaucoup aujourd'hui sur des processus de gestion de l'information publique, opère ici une redéfinition de l'acteur public en fournisseur de données normalisées à des tiers»³³. Les plateformes et les portails métropolitains de l'opendata fonctionnent alors comme un acteur central à la fois « capacitant »³⁴ ses utilisateurs et en délimitant les compétences. Ainsi tendue vers des nécessités internes, quasi de services, liées à l'injonction d'ouverture celle de la réutilisation ne peut être pensée première ou centrale. Les collectivités territoriales abordent la réutilisation « à partir de leur propre héritage institutionnel et à partir de leur posture « top down » »³⁵ et se préoccupent davantage de la conformité des démarches de l'ouverture des données aux cadres législatifs français et européen que des besoins et attentes des réutilisateurs. Et effectivement, les portails métropolitains de l'opendata en France restent pour les réutilisateurs professionnels des « magasins de pièces détachées » (Entretien avec un développeur). De plus, la réutilisation est complexe et réclame des compétences spécifiques et du temps car les données doivent être travaillées, croisées, nettoyées, extraites, traitées... « Un ré-utilisateur idéal, c'est quelqu'un qui a du temps, qui ne va pas faire un produit commercial dans l'année, mais qui a deux-trois années devant lui pour faire quelque chose » (Entretien avec un développeur). Beaucoup de réutilisations (prototypes et applications) restent d'ailleurs invisibles car ils disparaissent avant même d'être identifiés faute de modèle économique viable.

Rendre les données disponibles ne suffit pas pour générer la réutilisation et ce n'est sans doute même pas la première étape car c'est un processus complexe qui concerne plusieurs communautés professionnelles ayant différentes finalités stratégiques, économiques et éthiques. Cependant, une chaîne commune de réutilisation commune se construit autour de quatre étapes clés chez les réutilisateurs : la recherche de la donnée, son interprétation, sa transformation et son intégration. En effet, tous les réutilisateurs sont concernés par la recherche, localisation et téléchargement des jeux de données. Une fois la donnée trouvée et téléchargée, elle est cependant rarement utilisable en tant que telle. Il s'agit de la travailler afin de s'assurer de sa qualité, fiabilité et exploitabilité (fréquence de mise à jour, structure des données, exhaustivité). Dans tous les cas, la donnée doit être (re)-contextualisée, car extraite de sa source, elle en perd en partie son sens. En fonction des objectifs, des algorithmes utilisés et des savoir-faire internes, la donnée subit ensuite différents transformations et traitements. Il peut s'agir de nettoyer et de formater les données initiales dans des formats spécifiques, de consolider, redresser, agréger et enrichir les données en rajoutant par exemple des métadonnées ou en croisant les données issues de différentes sources. Certains professionnels intègrent les données de façon automatique, d'autres manuellement en les déposant dans un intranet ou un catalogue de données.

Pour que les données soient (ré)utilisées et pour que les réutilisateurs puissent les manipuler et transformer en informations ou services utiles, il faut que leurs descriptions, leurs formats et leurs métadonnées soient compatibles avec les besoins et les savoir-faire des réutilisateurs qui n'ont que très peu de dénominateurs communs. En effet, les écosystèmes spécifiques et structurés conditionnent les caractéristiques des données ouvertes utilisées qui doivent pouvoir s'intégrer dans les API, Intranet ou plateformes déjà en présence. Agir sur ces couches techniques de format et de standardisation est une gageure encore trop lourde pour être emportée dans les mois

³³ J. -B. LE CORF, « Management de l'information publique et innovation numérique de services urbains : l'intelligence territoriale en perspective » in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. BROUDOUX, G. CHARTRON (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 110.

³⁴ P. FALZON, "Enabling safety: issues in design and continuous design", *Cognition, Technology and Work*, n°1 (10)/2007, p.7.

³⁵ J.-M. NOYER, M. CARMES, « Le mouvement « Open data » et les intelligences collectives », in *Les débats du numérique*, M. CARMES, J.-M. NOYER (dir.), coll. « Territoires numériques », Paris, Presses des Mines, 2013, p. 151.

à venir, alors que poser les règles de jeu de l'ouverture semble un challenge plus abordable et tout autant nécessaire.

B) Poser un cadre de fonctionnement, si possible pérenne

De façon atypique à l'échelle nationale, les jeux de données du portail métropolitain lyonnais sont publiés sous trois types de licence, chacune adaptée à une certaine catégorie de données. Tout d'abord, la majorité des données sont publiées sous une Licence Ouverte, créée par la mission Etalab en 2011 et compatible avec les licences Open Government Licence du Royaume-Uni, Open Data Commons Attribution de l'Open Knowledge Foundation et Creative Commons Attribution 2.0 de Creative Commons³⁶. Cette licence ne pose aucune condition de réutilisation. Ensuite, le Grand Lyon propose une licence engagée. Gratuite, elle impose une déclaration préalable. Enfin, la licence associée introduit la possibilité d'une redevance modulable en fonction du succès commercial du produit ou du service créé. Véritable moteur complémentaire du développement économique auquel le PMOD s'associe, la licence associée veut « garantir un écosystème concurrentiel équitable, en évitant la formation de monopoles. Pour cela, cette licence outre l'authentification du ré-utilisateur, introduit la possibilité d'une redevance en fonction d'un seuil élevé du taux de pénétration du produit ou du service créé. En cas de position tendant à devenir monopolistique la redevance est élevée. Cette licence n'intègre aucun frais fixes afin d'éviter toute barrière à l'entrée pénalisante pour les petits acteurs. Elle vise à protéger les PME innovantes et acteurs de tailles modestes et favoriser leur développement. »³⁶.

Ces deux dernières licences visent clairement à promouvoir la réutilisation commerciale des données, mais la déclaration obligatoire qu'elles réclament permet également à la Métropole de mieux connaître ses ré-utilisateurs, de répondre à leurs besoins et de mettre en place une communauté de ré-utilisateurs. Ainsi, affine-t-elle et ajuste-t-elle son offre car elle cible mieux ses interlocuteurs, et évolue progressivement. À la création du PMOD lyonnais, les développeurs étaient la cible privilégiée avec un objectif manifeste de contribution au développement économique du territoire, à la croissance et la création de l'emploi dans une logique d'offre et de mise à disposition que nous avons déjà évoquée. Depuis 2015, l'idée de permettre aux citoyens de se réapproprier leurs propres données territoriales, de proximité, via le portail « a fait son chemin » (Entretien avec un gestionnaire de PMOD).

D'ailleurs, à ce jour, le mot de bienvenue du portail s'adresse à un large public : « Pour accélérer l'innovation et encourager la participation citoyenne, la Métropole de Lyon ouvre les données de référence de son territoire.... À vous de les découvrir et de les exploiter ! ». Cependant, une analyse plus complète que nous avons déjà présentée (Dymytrova, Hare, 2016) distingue clairement deux voies et cibles. D'une part, celle de l'innovation, de la création de services, porteuses de valeur ajoutée et de développement économique du territoire, pensée aux mains des professionnels de la data, du territoire, des acteurs économiques locaux ; de l'autre, celle de la citoyenneté et du partage collaboratif, porteuse de convivialité, de liens sociaux, proche d'une urbanité numérique qui rendrait la ville plus ou mieux habitable³⁷. Ce choix de positionnement se matérialise dans la rubrique « Utiliser la plateforme Data du Grand Lyon » qui comporte pour la réutilisation, deux entrées distinguant le « Coin de développeurs » et le « Coin des débutants », avec tous les sous-entendus qu'elles véhiculent.

Outil central de la panoplie OD développée par la Métropole de Lyon, ce portail est le guichet unique d'accès aux données d'un territoire car il centralise les données locales de la façon la plus complète possible en cherchant à y agréger toutes les données métropolitaines, même celles qui ne sont pas soumises à l'injonction publique. Mais il prolonge les choix politiques locaux en

³⁶ <http://data.grandlyon.com/connaître-nos-licences/> (consulté le 3 février 2017).

³⁷ Une ville habitable se laisse approprier et coloniser par ses habitants en tendant à devenir ville durable, contrairement aux mégapoles contemporaines qui deviennent inhumaines. Cf. Vanessa Timmer et Nola Kate Seymoar, *La ville habitable*, International Centre for Sustainable Cities, Vancouver, 2005

cherchant à assurer un lieu d'intermédiation entre les producteurs de données, les collectivités territoriales et les réutilisateurs de données³⁸, qu'ils soient de simples citoyens ou pas. La position tenue par le PMOD de Lyon est donc d'assurer la fonction centrale de l'entremise qui correspond bien avec la mission de la Métropole. Cependant, les moyens qu'elle se donne pour la tenir : faire du courtage informationnel entre les parties productrices et réutilisatrices des données, semblent aller à l'encontre de celle-ci. Double mouvement paradoxal qui fragilise particulièrement l'acteur central, la Métropole³⁹.

§ 3 – LE CADRE SOCIO-TECHNIQUE DE LA RÉUTILISATION DES DONNÉES

Le cadre socio-technique, tel que conçu et présenté par Flichy⁴⁰ surgit des « différentes histoires parallèles qui ne sont pas liées entre elles. C'est l'histoire antérieure à la constitution du cadre socio-technique qui reprend les histoires parallèles des mondes sociaux concernés ». Dans le contexte des PMOD, chaque réutilisateur est effectivement porteur d'une histoire parallèle, liée à son propre écosystème et à ses propres finalités et le cadre socio-technique ne peut se construire que dans l'action commune que le portail matérialise et symbolise. Ainsi tenter de cerner le réutilisateur et en comprendre l'histoire parallèle en termes de besoins et d'horizons d'attentes⁴¹ constitue un premier pas vers la constitution du cadre socio-technique de la réutilisation des données.

A) Tenter de cerner le ré-utilisateur

La plupart des documents de communication des collectivités font référence à trois types d'acteurs de l'Open data : les producteurs, les intermédiaires et les utilisateurs finaux comme le soulignait déjà Chignard dès 2012⁴². Sans invalider cette typologie notre enquête a révélé une variété de profils des réutilisateurs des données ouvertes qui pose problème car elle regroupe sous un intitulé théoriquement fédérateur, des métiers, des compétences et des positions hiérarchiques différents. Notre panel de réutilisateurs se compose en effet de producteurs de données, de gestionnaires de projets OD, de coordinateurs des projets basés sur l'utilisation des données ouvertes, de data-analystes ou journalistes, de développeurs et de promoteurs de données ouvertes. Sur le terrain, la catégorie structurante d'« usager intermédiaire » n'a pas de réalité claire ou circonscrite car un acteur peut se situer à l'interface de plusieurs positionnements, par exemple, les développeurs ou les bureaux d'étude qui réutilisent les données ouvertes pour générer leurs propres données et les réexporter par la suite auprès collectivités ou, dans d'autres cas, pour les vendre à leurs clients dans des applications clef en main. Selon leurs métiers et fonctions, les réutilisateurs interviewés utilisent les jeux de données pour réaliser des bases de données, des applications mobiles et web, pour faire de la visualisation ou rédiger des rapports ou conduire des analyses.

Autrement dit, chaque acteur conserve ses spécificités et ses particularités dans sa manière d'utiliser la donnée, ce qui oblige ces professionnels à tenir une double injonction impossible : ils ont besoin de données le plus brutes possibles pour les intégrer ou les traiter plus facilement alors

³⁸ V. LARROCHE, M. VILA, « Urban Data et stratégies dans le secteur des services : Le cas de la métropole lyonnaise », in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. BROUDOUX, G. CHARTRON (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 183.

³⁹ F. PAQUIENSÉGUY, « Les portails Open Data au prisme du courtage informationnel : qu'est-ce qui se joue pour les Métropoles ? », in *Open data : Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, F. PAQUIENSÉGUY (dir.), Paris, Editions des Archives contemporaines, 2016.

⁴⁰ P. FLICHY, *L'innovation technique. Récents développements en sciences sociales. Vers une nouvelle théorie de l'innovation*, Paris, La Découverte, 1995, p.224.

⁴¹ R. KOSELLEK, *Vergangene Zukunft. Zür Semantik geschichtlicher Zeiten*, Francfort, Suhrkamp, 1979.

⁴² S. CHIGNARD, *Open data : Comprendre l'ouverture des données publiques*, Paris, FYP Editions, 2012.

que sans son contexte, une donnée reste inutilisable. Mais malgré la diversité des profils de réutilisateurs interviewés, leurs besoins et les difficultés qu'ils rencontrent dans leur utilisation des données ouvertes ont de points de convergence.

B) Horizons d'attente des réutilisateurs professionnels

Fortement ancrées dans un territoire ou liées à un producteur, les données ouvertes sont éclatées entre différents portails et plateformes. Les données diffusées par chaque collectivité sont publiées sur un portail et une plateforme spécifique. Les mutualisations des publications se font en fonction des découpages administratifs, par exemple, à l'échelle d'une métropole ou d'une région. Certaines communes diffusent leurs données à la fois leur propre portail et celui de data.gouv.fr. Seules les petites collectivités territoriales publient leurs données uniquement sur le portail national. Mais les données ouvertes sont également hébergées sur des portails spécialisés d'administrations ou institutions comme l'Insee ou encore sur ceux des entreprises privées publiant leurs données comme data.sncf.com. D'une manière générale si la continuité et la prédictibilité de l'évolution de la donnée dans le temps sont très importantes pour tous les réutilisateurs leurs difficultés sont moins techniques qu'organisationnelles ou décisionnelles, à cause du caractère politique et stratégique de l'ouverture des données pour les acteurs territoriaux.

En effet, dans ce contexte très peu structuré et assez hétéroclite, plusieurs interviewés aspirent à la mise en place d'un outil commun des données ouvertes au niveau national, voire européen : « Une plateforme par ville ou par transporteur c'est un peu extrême » (Entretien avec un développeur). De plus chaque PMOD définit lui-même les thématiques des jeux de données à publier et les données qui seront, ou pas, publiées dans le respect des directives et lois. Ajoutons à cela le fait que le type de données disponibles varie en fonction des compétences spécifiques de chaque collectivité. L'absence de la donnée recherchée reste aujourd'hui un problème récurrent. Certaines données sont « prises en otage » en absence d'une obligation de publication (Entretien avec un gestionnaire de PMOD) : « En fait, dans l'Open data, c'est un peu la métaphore du placard où il n'y a pas tout ce qu'on veut pour faire le plat qu'on voudrait » (Entretien avec un développeur). Au-delà d'un simple accès à la donnée, beaucoup de réutilisateurs soulignent la nécessité de « comprendre le point de vue du producteur de la donnée : la donnée, ce n'est pas suffisant, il faut aussi les moyens de l'interpréter et donc des spécifications sur ces données et éventuellement des rapports sur l'utilité de ces données pour pouvoir vraiment la mettre en contexte » (Entretien avec un data-analyste). Toutefois, certains développeurs utilisent les données sans forcément se poser de questions sur la méthodologie du traitement de la donnée avant sa publication : « A partir du moment où je peux me dire que ces données sont relativement bonnes à la rigueur je m'en moque... Ils font leur travail. Pour moi, c'est une sorte de boîte noire, je récupère les données, pour l'instant cela me va. » (Entretien avec un développeur).

La compréhension des données est surtout possible grâce aux métadonnées. Or, à l'heure actuelle, les descriptions de données publiées (termes, mots-clés) et les métadonnées relatives à la création, aux mises à jour, aux utilisations sont disparates tant du point de vue sémantique que structurel entre les divers PMOD. Malgré l'existence de normes et de guides de bonnes pratiques que nous avons évoqués, chaque producteur fournit ses données dans son format spécifique sans se soucier des questions de l'exploitabilité. Par exemple, « il existe en France 7 fournisseurs des données en temps réel liés aux transports collectifs et 7 formats différents d'API » (Entretien avec un coordinateur de projet OD). Les réutilisateurs professionnels ne peuvent pas automatiser le traitement des données hétérogènes ce qui les oblige à un travail plus lourd d'agrégation manuelle et de consolidation afin de pouvoir répliquer leurs applications à l'échelle de plusieurs collectivités territoriales. L'accès aux données via les API freine leur emploi car « une API est une surcouche qui donne des services en plus qui sont les bienvenus, mais du coup il y a des restrictions...c'est bien pour un développeur qui veut tester des usages, mais si l'on veut faire une

entreprise derrière, il faut faire attention parce qu'il peut y avoir des restrictions par rapport au volume de données utilisées » (Entretien avec un développeur).

De plus, les formats des données disponibles viennent souvent des DSI des collectivités qui choisissent des formats plus simples ou des formats métiers qui ne sont pas toujours portés par des professionnels d'autres secteurs. En même temps, le choix des formats de données reste un choix politique et d'éthique professionnelle : « si tout le monde utilise GTFS, Google bouffe tout... » (Entretien avec un gestionnaire de PMOD).

Dans ce contexte, le souhait récurrent des réutilisateurs porte sur la structuration et la codification homogène et systématique des données des différents PMOD. La mise en place d'un référentiel multimodal de données permettrait d'harmoniser des données hétérogènes en ménageant l'interopérabilité afin d'avoir une plateforme structurée malgré la diversité hébergée. Le portail national data.gouv.fr n'est qu'en partie considéré comme un outil commun et pour les développeurs, le meilleur exemple actuel de véritable base commune des données ouvertes s'appelle *OpenStreetMap*.

La réutilisation des données ouvertes interroge ainsi les modes de gouvernance, les modèles économiques et les façons de penser le politique et l'espace public dans deux registres structurant. D'abord celui de l'accès aux données, informations et savoirs qui relève de l'ouverture elle-même et ensuite dans celui de la production-circulation des savoirs et services sous de nouvelles formes et conditions qui lui, relève de la réutilisation. Les deux sont totalement interdépendants car « des métadonnées aux fabrications (jusqu'aux générations semi-automatiques ou automatiques) d'ontologies et à leur interopérabilité partielle, ce qui se trame et est en jeu, c'est non seulement la création de valeur..., mais aussi différentes façons de concevoir le travail intellectuel, de concevoir et pratiquer les problèmes liés aux sociétés ouvertes et démocratiques. »⁴³.

Répondant aux injonctions de la transparence, de l'innovation et de la participation citoyenne, la démarche de l'ouverture des données croise souvent les préconisations des promoteurs du concept de « ville intelligente », imaginée comme une plateforme d'innovation ouverte, basée sur les nouveaux services urbains. À la différence des services publics traditionnellement centralisés et descendants, les services urbains d'une smart-city reposeraient sur une implication des usagers « à la construction ou à la fourniture d'une partie du service », tout en considérant les usagers « à la fois [comme] les unités d'observations, les actants et les bénéficiaires. »⁴⁴.

Autrement dit, la réutilisation, par des professionnels, des pro-am - et des citoyens marquera la réussite de l'ouverture des données et sa promesse politique et utopique d'une réutilisation par tous.

BIBLIOGRAPHIE

S. CHIGNARD, *Open data : Comprendre l'ouverture des données publiques*, Paris, FYP Editions, 2012.

V. DYMYTROVA, I. HARE, « La mise en scène visuelle de l'Open data : enjeux sémiotiques, communicationnels et politiques », communication au colloque *Les données urbaines : quelles pratiques ? quels savoirs ?*, Lyon, 13-14 décembre 2016.

⁴³ J.-M. NOYER, M. CARMES, « Le mouvement « Open data » et les intelligences collectives », in *Les débats du numérique*, M. CARMES, J.-M. NOYER (dir.), coll. « Territoires numériques », Paris, Presses des Mines, 2013, p. 157.

⁴⁴ B. MARZLOFF, *Le 5^e écran, les médias urbains dans la ville 2.0*, Paris, FYP Editions, 2009, p.66.

- F. IBEKWÉ-SANJUAN, F. PAQUIENSÉGUY, « Open, Big, Collaboration : trois utopies de l'innovation au 21e siècle », in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. Broudoux, G. Chartron (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 15.
- P. FALZON, « Enabling safety: issues in design and continuous design », *Cognition, Technology and Work*, n°1 (10)/2007, p.7.
- P. FLICHY, *L'innovation technique. Récents développements en sciences sociales. Vers une nouvelle théorie de l'innovation*, Paris, La Découverte, 1995.
- M. FRANK, J. WALKER, « User centred methods for measuring the quality of open data », *The Journal of Community Informatics*, n°12/2016, p. 47.
- R. KITCHIN, *The Data Revolution: Big data, Open data, data infrastructures and their consequences*, London, Sage, 2014.
- V. LARROCHE, M. VILA, « Urban Data et stratégies dans le secteur des services : Le cas de la métropole lyonnaise », in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. BROUDOUX, G. CHARTRON (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 183.
- J. -B. LE CORF, « Management de l'information publique et innovation numérique de services urbains : l'intelligence territoriale en perspective » in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. BROUDOUX, G. CHARTRON (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 97.
- R. KOSELLEK, *Vergangene Zukunft. Zur Semantik geschichtlicher Zeiten*, Francfort, Suhrkamp, 1979.
- B. MARZLOFF, *Le 5^e écran, les médias urbains dans la ville 2.0*, Paris, FYP Editions, 2009.
- D. A., NORMAN, S.W. DRAPER (Eds.). *User-centered system design: New perspectives on human-computer interaction*, Hillsdale, NJ, Lawrence Erlbaum, 1986.
- J.-M. NOYER, M. CARMES, « Le mouvement « Open data » et les intelligences collectives », in *Les débats du numérique*, M. CARMES, J.-M. NOYER (dir.), coll. « Territoires numériques », Paris, Presses des Mines, 2013, p. 137.
- F. PAQUIENSÉGUY, « A qui profite le chiffre? Les portails métropolitains Open Data », in *Datalogie. Formes et imaginaires du numérique*, O. AVENATI, P.-A, CHARDEL (dir.), Editions Loco en coproduction avec l'ESAD de Reims, 2016, p. 126.
- F. PAQUIENSÉGUY, « Smart city & open data : à qui profitent les données ouvertes ? », *CIST2016 - En quête de territoire(s) ?*, Mars 2016, Grenoble, France. Proceedings du 3e colloque international du CIST, p.351. <https://hal.archives-ouvertes.fr/hal-01353651/>.
- F. PAQUIENSÉGUY (dir.), *Open data : Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, Paris, Editions des Archives contemporaines, 2016.
- F. PAQUIENSÉGUY, « Les portails Open Data au prisme du courtage informationnel : qu'est-ce qui se joue pour les Métropoles ? », in *Open data : Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, F. PAQUIENSÉGUY (dir.), Paris, Editions des Archives contemporaines, 2016.
- S. TURKI, M. FOULONNEAU, « Valorisation des données ouvertes : acteurs, enjeux et modèles d'affaires », in *Big data - Open data: Quelles valeurs? Quels enjeux?*, E. BROUDOUX, G. CHARTRON (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 113.
- R. K. YIN, *Case Study Research. Designs and Methods*, Newbury Park, Sage, 2009.

