L'Open Data métropolitain, en voie de structuration ?

Françoise Paquienséguy
Sciences Po Lyon / LabEx IMU - Elico

Résumé : Les métropoles développent depuis maintenant plus de dix ans leurs portails Open Data afin de respecter l'injonction européenne d'ouverture des données publiques liées au territoire. Par-delà la diversité des expériences, des éléments de structuration organisationnelle apparaissent, moins spécifiques et particuliers que les portails eux-mêmes, souvent en lien avec l'idéologie d'une ville intelligente fondée sur les données. Ce chapitre analyse les stratégies d'acteurs à l'œuvre pour dégager ces constantes et faire la démonstration d'un processus de structuration de l'Open Data, pensé comme une filière de l'industrie de la donnée. La figure du réutilisateur y est centrale.

Mots-clés : Portails métropolitains Open Data, business model, structuration, idéologie, filière de l'Open Data, réutilisation

1 Introduction

Les premiers éléments liés à l'Open Data remontent maintenant à plus de dix ans, que l'on convoque le *Freedoom of information Act* américain ou la directive européenne Inspire. Sur la période, ce chapitre cherche à faire le bilan des actions des métropoles françaises au travers de leurs portails Open Data en considérant l'idéologie de l'ouverture comme un des cadres structurants de l'Open Data. En essence, il tente de répondre à la question qui lui sert de titre. En effet, malgré toute la diversité des cas, des portails, des objectifs et des données ouvertes, elles-mêmes souvent qualifiées d'hétérogènes, la décade d'expérimentations immédiates, de structurations et de recompositions des acteurs, des services et des moyens a conduit à une certaine stabilisation, voire structuration de l'écosystème de l'Open Data métropolitain en France qui permet de dégager à la fois des points clefs communs et les prémices d'un modèle économique.

2 Revenir à l'idéologie initiale de l'ouverture des données

Revenir sur l'idéologie qui a porté et entraîné l'ouverture des données paraît indispensable pour comprendre l'importance du lien entre *Open Data* et *Open Government*, tant ces notions sont pensées ensemble dans cette idéologie. Plusieurs textes, dont ceux de Tom O'Reilly (2010) principalement, soulignent que les données ouvertes constituent le socle de la transformation de l'action publique d'une part, et de l'autre que l'*Open Government* n'est pas simplement une « *gouvernementalité algorithmique* » (Rouvroy et Berns : 2013), mais bien une autre forme de démocratie, initialement présentée comme disruptive.

C'est à partir des fonctions de l'idéologie définies par Ricœur (1983) que nous saisirons celle de l'ouverture des données pour comprendre les postulats qu'elle a formulés et les objectifs qu'elle a dessinés. L'idéologie, ses rhétoriques, ses faits marquants et ses défenseurs ont marqué l'ouverture des données qui dès le départ est porteuse de deux tendances, l'une à visée plutôt technologique, ancienne, issue de la cybernétique (Ibekwé et Paquienséguy : 2015) visant à produire normes, formats et l'autre, née juste après le siècle, visant à proposer d'autres formes d'action publique et de gouvernance. En effet, avec l'utopie, l'idéologie est une des deux formes d'expression de l'imaginaire social (Ricœur : 1983, 53-54), dont trois usages majeurs éclairent celle de l'ouverture des données.

Premièrement, selon Paul Ricœur, une idéologie sert à la dissimulation ou à la distorsion de la réalité sociale, de la praxis (*Ibid.*, 54). De façon frappante, la fonction de distorsion portée par l'idéologie de l'ouverture des données au début du XXIe siècle est tout d'abord de poser la data comme élément premier et central, par lequel l'information, la connaissance et la prédiction s'introduiront ensuite dans la praxis, et ici la prédiction s'entend plus comme un outil d'aide à la décision que comme un processus de profilage. L'idée centrale, fondatrice de cette idéologie est « *la mise à disposition de données numériques non confidentielles, pour tous, sans restriction sur le droit d'accès et de réutilisation* » et les critères essentiels de l'ouverture sont « *la disponibilité, la réutilisation, la distribution et la participation universelle* » selon la définition, une des premières, donnée en 2005, par l'*Open Knowledge Foundation*. Le seul nom de la fondation formalise déjà ce passage de la data à la connaissance, la seconde dépendant de la première. La distorsion se situe bien là : une connaissance fondée sur des données ouvertes, accessibles, partagées. D'ailleurs cette association prône l'ouverture de tous les types de contenus assis sur des données puisqu'elle s'inscrit dans un concept plus large d'*Open Knowledge* [1] où la connaissance correspond à l'œuvre (film, musique, article scientifique ou data). Si, ici, l'idéologie de l'Open Data vérifie, selon les principes de P. Ricœur, les fonctions de dissimulation et de distorsion, le dernier ouvrage du fondateur de l'*Open Knowledge Foundation*, Rufus Pollock, professeur d'économie à Cambridge, précise clairement que ces fonctions, dans le cadre de l'idéologie de l'ouverture, conduisent soit à « une dictature numérique aux mains des géants de la donnée [2], soit à une démocratie du numérique aux mains des citoyens » (Pollock : 2018, 8-9). Autrement dit, la fonction de distorsion de l'idéologie de l'ou-

1. Voir le site de la section française https ://fr.okfn.org/, repéré le 14 décembre 2019
2. GAFAM et BATX principalement.

verture articule fermement data, information et connaissance plaçant bien la data en début des chaînes de médiation des données ; alors que la fonction de dissimulation en occulte l'objectif final et l'aboutissement, entre dictature et démocratie numérique, entre liberté et aliénation de l'individu au regard des *data compagnies* et de leurs algorithmes.

Deuxièmement, P. Ricœur considère l'idéologie comme une « *rhétorique du discours public mise au service du processus de légitimation de l'autorité* » (*Ibid.*, 56) et affirme donc son pouvoir à légitimer même s'il l'analyse comme seconde au regard de la fonction d'intégration, englobante. Le processus d'ouverture de données se nourrit des rhétoriques de l'action publique, tout d'abord au niveau européen puisque dès 2003 le Parlement Européen cadre la réutilisation des informations du secteur public dans la directive européenne 2003/98/CE (dans la lignée de la loi CADA de 1978), puis confirme l'ouverture des données dans la fameuse directive Inspire de 2007. Ce niveau européen va à la fois délimiter et neutraliser le débat, les pays membres se limitant à appliquer les recommandations de la directive. Ensuite, le discours national viendra assurer la jonction avec le citoyen selon ses propres orientations. Pour la France, l'ouverture des données s'inscrit dans « *une tradition de transparence démocratique et de partage des informations détenues par la puissance publique* ». Tradition qui l'a conduite en 2016-2017 à assurer pour la première fois la présidence de *l'Open Government Partnership*[3] dont elle est membre depuis 2014, suite à la mise en place d'Etalab en 2011. Fondé la même année, ce partenariat, initié par Hillary Clinton[4] s'inscrit dans la foulée du plan *Open Government* de l'administration Obama[5] qui cherche à préserver la vision disruptive évoquée plus haut.

Ainsi se dessinent déjà deux visées idéologiques différentes de l'ouverture des données qui se complètent et génèrent une synergie qui pousse à l'action.

Il y a d'une part la rhétorique publique nationale, amplifiée à l'international par des organismes ou leaders qui soutiennent l'ouverture selon une idéologie de liberté, de transparence et de participation comme Tom O'Reilly, dans son texte de 2010 *Open governement*, dans lequel l'ouverture des données doit être « *an unprecedented level of openness in government* ». D'ailleurs, le leadership de la France à la tête du partenariat pour un gouvernement ouvert s'inspire de cette position et la phrase clé de ce mandat, assumé par François Hollande, montre bien le caractère idéologique de l'ouverture des données : « *bâtir un gouvernement ouvert, c'est œuvrer pour le renouveau de la démocratie et de l'action publique* ». Cette vision de l'Open Data promeut une action publique horizontale dans laquelle la participation citoyenne passe par des plateformes collaboratives.[6] Pour la France, elle s'appuie sur la mise en place d'Etalab en 2011, sur la multiplication des portails de données ouvertes issus de collectivités territoriales, sur la promulgation de la loi pour une République Numérique en 2017, et dernièrement en 2018, sur le plan État Plateforme 100% Numérique. Ce dernier fait du numérique

3. http ://www.opengovpartnership.org, repéré le 29 juin 2019
4. https ://www.state.gov/documents/organization/164473.pdf, repéré le 17 juillet 2019
5. https ://fas.org/sgp/crs/secrecy/R41361.pdf, repéré le 14 juillet 2019
6. https ://laspic.hypotheses.org/1896, repéré le 30 juin 2019

et de l'accès aux données, l'instrument de transformation de l'État et de l'évolution de la démocratie [7], promesse de campagne d'Emmanuel Macron.

Mais d'autre part, il y a la rhétorique d'acteurs associatifs à but non lucratif et non partisan, indépendants des gouvernements, comme *l'Opening Knowledge Foundation* ou la *Sunlight Foundation*, pour qui l'ouverture promeut avant tout l'accès citoyen à l'information gouvernementale et aux données publiques. Pour se développer et être facilité, voire équitable, cet accès suppose principalement de produire des normes techniques (formats) et juridiques (licences), et de viser la réutilisation des données, ici comprise comme leur appropriation par le citoyen au bénéfice de son action. La première fondation œuvre donc tout particulièrement à la création de licences (comme la licence ODbL) et plaide pour le libre dans ses différentes formes (contenus, savoirs, outils) ; alors que la seconde développe des technologies citoyennes (dites civic tech) afin de faciliter l'action citoyenne dans l'action publique, du point de vue citoyen, on l'aura bien compris, complété par des acteurs issus de la société civile comme la Fing ou Regards Citoyens évoqués *infra*. Ainsi, l'écosystème de l'Open Data s'en trouve-t-il complet entre action publique, contrepoids, adjuvants associatifs de rang international et acteurs de terrain de l'ouverture, au premier rang desquels les métropoles, puisqu'il s'agit par-là d'accompagner « *la modernisation de l'action publique territoriale et l'affirmation des métropoles* » [8]. Par sa structuration, le processus d'ouverture formalise et légitime l'action publique dans une forme contemporaine de démocratie qui, soit valorise l'action citoyenne à travers des plateformes nationales (comme data.gouv.fr), soit promeut une gestion citoyenne de la cité qui combine collaboration et prédiction, autrement dit qui flirte avec les idéaux de la *Smart City*.

Enfin, l'idéologie a pour troisième fonction de « *diffuser la conviction que ces événements fondateurs sont constitutifs de la mémoire sociale et, à travers elle, de l'identité même de la communauté* » (Ricœur : 1983, 58). Rappelons-nous qu'initialement, le concept d'ouverture est issu de la communauté scientifique car « *c'est une pratique ancienne qui fait partie de ses socles. Et l'idéal serait sans doute pour elle que toutes les données ayant servi à réaliser une étude soient accessibles dans un format ouvert afin de permettre à d'autres scientifiques de s'en saisir* » (Ibekwé et Paquienséguy : 2015) ; autrement dit, l'Open Science. Sur le même principe de propagation, le lien se fait immédiatement avec des déclarations et actions publiques d'envergure qui accompagnent à la fois l'ouverture et l'exploitation des données. Depuis le rapport *Villani* commandé en 2016 par Emmanuel Macron, alors ministre, et Axelle Lemaire, alors secrétaire d'État chargée du numérique, le chef de l'état confirmait récemment une ouverture massive et généralisée des données d'intérêt général : « *Nous continuerons de renforcer et d'élargir le service public de la donnée, c'est-à-dire la liste des bases publiques ouvertes et maintenues avec un très haut niveau de qualité, de fraîcheur de la donnée et de disponibilité du service* » [9]. De façon plus large, les marqueurs rhétoriques et discursifs sont trop nombreux (Dymytrova et Hare : 2016) dans la sphère publique pour être mentionnés ici, mais ils accompagnent la France dans un proces-

7. https://www.gouvernement.fr/action/le-numerique-instrument-de-la-transformation-de-l-etat, repéré le 30 juin 2019

8. Loi du 27 janvier 2014

9. Discours du 29 mars 2018, prononcé par Emmanuel Macron à l'occasion de son allocution sur *l'intelligence artificielle au service de l'humain*. Disponible ici : , repéré le 29 juin 2019

sus d'ouverture. Quelques exemples sont d'ailleurs emblématiques comme l'*Open Data Lab* national et ses dérivés régionaux qui produisent des *Guides pratiques de l'ouverture* ; comme lorsque E. Macron entend passer aux « *travaux pratiques* » et que la secrétaire d'État chargée du Numérique croit « *en notre modèle de services publics, en son avenir à l'âge numérique et en l'alliance entre les écosystèmes innovants et citoyens* ». Le décalage entre ancrage identitaire national et action de terrain reste flagrant, nous le traiterons ci-après, mais le registre discursif s'accroche aux valeurs républicaines : liberté de l'accès, égalité de l'accès, partage des données, expression de la citoyenneté, etc. « *L'Open Data est donc à chaque fois [...] mis au service d'une cause démocratique* » [10]. À tel point que depuis la mise en place d'Etalab et de data.gouv.fr, l'ouverture des données publiques est une obligation pour les communes de plus de 3500 habitants [11], avec la loi Lemaire pour une République numérique, elle représente une action démocratique, sous la conduite d'Etalab « *Chief Data Officer de l'État français* » depuis 2014 [12].

C'est pourquoi, pour retrouver une relative neutralité entre les deux, avant de voir comment ces idéaux seront revisités lors de leur mise en œuvre, nous nous baserons désormais tout au fil de cet ouvrage, sur la définition de l'ouverture des données, de l'Open Data telle que formulée, en 2010, par *Wikipédia* de façon stable et générique, autrement dit, la définition de la communauté qui soutient le mouvement d'ouverture :

> *l'Open Data consiste à donner accès avec un usage libre à des données numériques. Ces données sont diffusées de manière structurée selon une méthode et une licence ouverte garantissant son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière.* [13]

2.1 Les idéaux de l'ouverture des données revisités

« *Le récit dominant de l'ouverture des données publiques décrit un avenir radieux* », résume Daniel Kaplan de la Fing, « mais la réalité est souvent plus complexe », confirme Sandra Moatti, rédactrice en chef d'*Alternatives Économiques*. En effet, les premiers bilans témoignent surtout d'une profusion d'actions. À l'échelle nationale, après plus de dix ans d'investissement, l'Open Data se matérialise par le portail national, le nom de domaine data.gouv.fr, des licences de plus en plus utilisées, une offre de plateformes sur étagère vendues par Opendatasoft, l'obligation de mise à disposition des données pour les communes de plus de 3500 habitants, 146 portails Open Data communaux ou métropolitains [14] et de très nombreux événements ou actions de promotion et d'accompagnement de rang national comme la *Quinzaine des données ouvertes*, les *Trophées de l'Open Data*, sans oublier la feuille de route « *Open Data par défaut* » du gouvernement Philippe pour 2018-2020 dans le cadre de l'*Open Government Partnership*. Sans conteste, l'Open Data est une réalité de l'action publique, les outils de référence et le cadre législatif sont en place et la promotion d'une culture de

10. https ://laspic.hypotheses.org/1896, billet de Jonathan Chibois, repéré le 29 juin 2019
11. Soit 1858 communes sur 36681, Insee, 2010
12. https://www.Etalab.gouv.fr/qui-sommes-nous, repéré le 7 juillet 2019
13. https ://en.wikipedia.org/wiki/Open*data, repéré le 3 décembre 2019
14. https ://www.opendatasoft.fr/ressource-liste-portails-open-data-dans-le-monde, repéré le 9 juillet 2019

l'ouverture a fonctionné. Cependant, la mise en œuvre, *via* plateformes, portails, accès et réutilisation ne lui permet toujours pas de « *libérer tout son potentiel* » (Frédéric Charles – GreenSI) [15]. Les usages, le ou les modèles économiques, la structuration interne des services en charge de l'ouverture restent encore à trouver ou à stabiliser, tant les cas diffèrent au gré des villes, des métropoles et de leurs propres idéaux et objectifs. Il est illusoire de penser l'ouverture par les administrations et collectivités territoriales comme la dernière étape d'un processus générant des données et libérant de la valeur. Elle réclame tout un travail d'arrière-plan, toute une chaîne de traitement de la donnée qui impose de les nettoyer, de les formater, de les actualiser, d'en fournir les métadonnées, avant de les proposer à la réutilisation et de définir une chaîne de production de valeur adaptée.

Ainsi, les acteurs publics concernés, et tout particulièrement les villes et métropoles que nous avons étudiées pendant trois ans, ont-ils ouvert les données en fonction de leurs propres contextes, politique, économique et idéologique (Paquienséguy et Dymytrova : 2017). De fait, aucune ville de taille intermédiaire, aucune métropole même ne porte le même projet d'ouverture des données. Leurs visions de l'ouverture, servie ou empêchée, par les moyens mis à disposition (moyens qui nécessitent une réaffectation des impôts locaux et passent souvent par une restructuration de la direction des systèmes d'information) varient bien sûr, il en sera longuement question dans cet ouvrage et de fait, le modèle économique reste souvent singulier et évolutif (Paquienséguy : 2016) sur la base de

> *deux idées essentielles, [...] concomitantes. D'une part, l'ouverture des données est un processus lent car lourd dont les deux premières étapes (ensemencement et développement) rapportent quasiment autant de gains économiques qu'elles n'en coûtent. D'autre part les gains les plus rapidement visibles, et constamment croissants, ne sont pas des gains économiques, mais plutôt des externalités positives. Ce qui laisse à penser qu'une bonne partie des bénéfices de l'Open Data se situe hors-marché (Mahé et Ennassiri : 2014)*

en lien avec le déploiement des services de la ville intelligente.

C'est pourquoi l'Open Data contribue autant à la transformation de la métropole en relation étroite avec des partenariats, des consortiums, des réseaux d'acteurs d'où proviennent des bénéfices hors-marchés, au regard de sa propre concrétisation des idéaux de l'ouverture des données.

En effet, souvent rattachés à des services stratégiques pour le développement économique du territoire, ces portails sont la trace de changements forts, clairs et assumés dans les stratégies territoriales car ils concrétisent des intérêts multiples, issus d'acteurs de natures très différentes qui font le tissu urbain, sur la base d'objectifs et d'intéressement (Callon : 1986, 170-208). Fruits d'un consensus produit par un « *réseau hétérogène* » (Law et Callon : 1992, 21-52) d'acteurs implantés dans le territoire métropolitain, ils matérialisent les choix de la collectivité qui les a conçus. La visua-

15. https://www.zdnet.fr/blogs/green-si/redefinir-l-open-data-pour-liberer-tout-son-potentiel-39871229.htm, repéré le 30 novembre 2019

lisation, les publics cibles et les modalités d'accès aux données proposés témoignent alors, à des degrés variables, de l'intégration des citoyens et/ou des professionnels de la data, dont les données ouvertes (Paquienséguy et Dymytrova : 2018).

Nous avons analysé [16] ces enjeux dans un corpus de huit portails métropolitains : Londres, Amsterdam, Lyon, Berlin, Séoul, Chicago, Montréal et New-York, tous étudiés dans le cadre de l'ANR *OpenSensingCity* (Paquienséguy et Dymytrova : 2017) et détaillé plusieurs résultats éclairants.

En effet, les métropoles étudiées travaillent à la mise en place d'un « *dispositif d'intéressement* » (Callon :1986) selon des registres différents. Ce dispositif porte « *l'ensemble des actions* » (ici la conception de la plateforme, ses règles de fonctionnement et les modalités de gouvernance) « *par lesquelles une entité » (ici la métropole qui reste leader) « s'efforce d'imposer et de stabiliser l'identité des autres acteurs* » (ici les producteurs, les développeurs, les réutilisateurs et les professionnels de la data) qu'elle a « *définis par sa problématisation* » (ouvrir, partager les données pour les valoriser). L'écosystème métropolitain est donc à la fois générique comme nous venons de l'exprimer dans les parenthèses et particulier puisque les acteurs locaux dominants et la thématique identitaire de la ville (le design, l'innovation, la culture, la ville intelligente, la ville créative, etc.) vont orienter à la fois la production et la réutilisation des données en fonction des objectifs stratégiques de développement économique de la ville et de ses acteurs. En effet, « *l'intéressement est fondé sur une certaine interprétation de ce que sont et veulent les acteurs à enrôler et auxquels s'associer* » (*Ibid.*, 188).

Ainsi, les données métropolitaines profitent-elles à différentes catégories d'acteurs lesquels choisiront à la fois les données à produire, celles à ouvrir et celles à exploiter ; ils se définiront alors comme les premiers destinataires du portail Open Data métropolitain. Le façonnant à leurs besoins et objectifs, ils le caractérisent et l'éloignent plus ou moins des deux idéaux initiaux au profit d'une réalisation technique à visée citoyenne ou économique, ou stratégique en termes de développement du territoire. C'est la phase de concrétisation du projet et les alliances d'acteurs qui en font les caractéristiques, variables dans la durée car les portails de 2019 ne ressemblent pas forcément à ce qu'ils ont été quelques années auparavant, l'Open Data se réalise en cours de route, au gré d'interactions entre contexte national et tensions locales du territoire.

2.2 Une intersection sensible et stratégique pour les acteurs du territoire

Les principales métropoles contemporaines se trouvent soumises à deux tendances dominantes issues d'injonctions diverses. Parmi elles, l'injonction publique bien sûr mais aussi l'importance croissante de la métropolisation (Charles et Thouément : 2018), confortée en France par la réforme territoriale de ces dernières années qui leur offre des compétences singulières, et enfin, plus largement, par la concurrence

16. Ces éléments ont initialement été présentés dans le cadre d'une communication au Colloque International des Sciences du Territoire en 2016 à Grenoble, *cf* Paquienséguy F., *Smart City & Open Data : à qui profitent les données ouvertes ?* Colloque international des Sciences du Territoires 2016, Proceedings « En quête de territoire (s), Looking for territories », Grenoble, 17-18 mars 2016, p. 351-357

internationale avec les villes capitales. Si les tendances internationales ne peuvent être mises de côté pour étudier le développement des portails Open Data tant elles pèsent sur les stratégies et objectifs des métropoles qui s'en trouvent encore plus renforcés, nous n'en retiendrons qu'une, centrale et fondamentale ici : devenir une ville intelligente. La deuxième tendance lourde serait devenir une ville créative, par exemple en (sur)-valorisant le patrimoine culturel, matériel ou immatériel, la science et l'art dans la métropole, mais nous ne l'étudierons pas ici car elle ne porte pas de stratégies autonomes à propos de data puisqu'elle s'appuie généralement sur celles pensées pour la *Smart City*. Si les villes créatives sont labellisées par l'Unesco, les villes intelligentes sont quasiment autoproclamées, sous la bannière nationale de la French Tech, et à l'échelle de la France, peu de villes, Lyon et Rennes principalement, jouent sur les deux registres smart et créatif, alors qu'en Chine ou en Inde la double démarche est plus commune [17].

En France, la cohésion du projet métropolitain de *Smart City*, renforcée par l'action publique, comme nous l'avons écrit en début de chapitre, fait du portail Open Data un objet stratégique car lieu de convergence des acteurs du territoire. En effet, il doit servir leurs intérêts et leurs objectifs, sans lien fort ou direct avec les premières idéologies, parfois bien lointaines. Il doit offrir

> *des données pour atteindre des objectifs dans une démarche de changement et de gestion intégrée reposant en grande partie sur l'interopérabilité des systèmes ; les acteurs, au sens large (institutions, organismes publics et privés, entreprises, organismes bancaires, marchés, individus…), peuvent tenter d'influer sur cet avenir, le diriger à leur façon en construisant des projets répondant à leurs natures, leurs missions, leurs valeurs.* (Chopplet : 2018,73).

Le portail Open Data se trouve donc au cœur : 1. d'un processus de facilitation de la réutilisation des données, il faut le rappeler puisque c'est l'objectif officiel premier ; 2. des stratégies de développement et de promotion du territoire, donc de ses acteurs économiques les plus présents, les plus puissants ; 3. de la transition numérique des villes par et avec les données vers une ville intelligente, créative, et/ou durable.

Deux caractéristiques prégnantes témoignent du rôle fort et premier des données métropolitaines et de leur portail dans la transition numérique de la ville, et tout particulièrement dans le cas de la Métropole de Lyon. Notons déjà, premièrement le rôle central de la gouvernance locale, acteur complexe et puissant porteur de la dynamique générale en termes d'actions (reconversion de certains quartiers, équipements ou transports publics, infrastructures, etc.) ; deuxièmement, la double exploitation des données à la fois en tant que telles (données brutes) et transformées en application (création de valeur ajoutée servicielle), double registre qui vient conformer les objectifs de *Smart City*. Autrement dit, l'équilibre entre le soutien, les objectifs de la métropole d'une part et l'exploitation directe ou servicielle des données de l'autre.

17. https://www.hindustantimes.com/real-estate/why-smart-cities-need-art-and-culture-to-thrive/story-axqPb391LEDv31OHDZDFNN.html, repéré le 13 juillet 2019

C'est pourquoi d'ailleurs, la métropole se structure, et ne cesse de se restructurer pour compléter le dispositif général de la donnée incluant les fonctions de production, de traitement, d'exploitation et de diffusion ainsi que les acteurs qui s'y réfèrent ; dans le droit fil de la directive Inspire de 2007 qui pousse à constituer tout un écosystème de fourniture, d'accès et de traitement des données publiques qui sera défini par les choix métropolitains.

3 Des objectifs encastrés et des stratégies d'acteurs complexes

Cette restructuration peut sans doute s'interpréter comme un des signes de la multiplication des enjeux et des objectifs portées par les portails métropolitains Open Data qui cumulent et fédèrent des ambitions variées, du niveau européen comme du territoire finalement concerné. Les stratégies d'acteurs sont en fait au cœur de la problématique de l'Open Data. Si en essence, les mêmes catégories d'acteurs semblent persister d'un cas à un autre, sur le terrain, l'écosystème de chaque métropole diffère par ses objectifs et missions. Au final, seuls certains acteurs sont présents par principe dans tous les cas de figure étudiés : les acteurs publics directement désignés par l'injonction d'ouverture et ceux qu'ils ont eux-mêmes créés ou mis en place. Cependant, s'ils s'incarnent dans des administrations pérennes, ils sont conduits par des élus dont le positionnement pèse sur les choix d'action bien entendu ; ainsi la couleur politique et les orientations des présidents de métropoles et des maires de la ville éponyme doivent-elles être considérées comme un niveau supplémentaire de complexité et de changement, au fil des élections municipales. De même les compétences de la métropole suite à la réforme territoriale doivent-elles êtres prises en compte puisqu'elles différent de façon conséquente. La Métropole de Lyon, par exemple est à ce jour la seule métropole française à exercer sur son territoire les compétences d'une métropole et d'un département depuis janvier 2015.

3.1 Des éléments structurants

Ces portails, et leurs écosystèmes matérialisent plusieurs éléments forts des stratégies des acteurs à l'œuvre. En effet, dans un premier temps, nous remarquons que ces portails Open Data des plus grandes villes françaises et internationales forment un « *réseau hétérogène* » d'acteurs dont la diversité intervient de façon variée, qu'il s'agisse de taille comme par exemple entre la métropole prise comme collectivité territoriale et un collectif de quartier ; qu'il s'agisse de puissance financière comme par exemple entre les grands groupes industriels concernés par l'urbain (Decaux, Vinci, Véolia, Bouygues, Enedis, Eiffage, etc.) et les start-up locales ou les services métropolitains eux-mêmes ; qu'il s'agisse de compétences par exemple entre les développeurs, les *data scientists* et les producteurs de données. A ces quelques exemples de diversité devraient s'ajouter ceux fournis par l'histoire, les objectifs et les modalités de décision de chaque catégorie d'acteurs. L'ensemble des « *acteurs sociaux mobilisés par la conception* » (Akrich : 1992, 7) du portail Open Data métropolitain doit donc s'accorder et négocier en vue de la mise en œuvre du portail lui-même. Cette première étape, matérialisée par la première décade des portails Open Data métropolitains (2007-2017) a permis « *la construction de liens inédits entre des réseaux jusqu'alors*

disjoints : il est clair que le nouveau réseau ainsi constitué conduit à une déformation des réseaux précédents et à une recomposition des alliances » (*Ibid.*, p. 5) que nous avions jusque-là considérée comme la structuration de l'écosystème territorial assortie de la restructuration de l'acteur public métropolitain et de ses composantes. C'est sans doute dans cette première étape que réside la dimension innovante de l'Open Data métropolitain car pour faire exister ces portails, considérés comme objets consensuels à inventer, les acteurs locaux ont dû négocier, s'allier et penser le dispositif de l'ouverture, autrement dit, fédérer des réseaux antérieurs, jusque-là séparés et en générer de nouveaux, spécifiques à ce nouvel écosystème.

L'évolution ou le renforcement des alliances conduit alors à positionner différemment à la fois l'acteur public (la métropole) et le territoire qu'il représente. « *Les acteurs et les objets ne doivent être considérés que comme le résultat du processus d'innovation et non comme le commencement* » (*ibid.*, 7). Sans pour autant acter définitivement qu'un *optimum* soit atteint, cette première phase de la jeune histoire des portails Open Data a vu se mettre en place plusieurs éléments structurants de l'écosystème et la métropole se poser en coordinatrice et fédératrice du réseau d'acteurs, malgré sa complexité. Le premier objectif de production et d'agrégation des données semble donc franchi, non sans mal puisque la plupart des métropoles auront produit trois à quatre versions différentes de leur portail sur la période, cherchant en permanence son adéquation avec les objectifs, les modalités et les choix techniques opérés par le réseau d'acteurs.

Dans la durée, quelques éléments techniques apparaissent d'ailleurs comme des constantes et se retrouvent sur quasiment tous les portails étudiés, faisant peut-être office de standard ou normes, de façon (encore) empirique. C'est le cas par exemple :

- du catalogue des jeux de données au format DCAT (*Data CATalog vocabulary*) qui permet de catégoriser tous les jeux de données offerts par le portail et d'en avoir à la fois une lecture et une extraction facilitée pour les réutilisateurs professionnels,
- de la récurrence de la majorité des thèmes retenus pour catégoriser les données qui permet une exploitation systématique des jeux de données à une échelle supra-métropolitaine d'une part (à l'échelle régionale si on prend l'exemple de Lyon/Saint-Etienne) ou nationale d'autre part (si on prend l'exemple de certaines applications comme *To Good to go* ou *Park & Go !*) ; sont particulièrement génériques à ce jour : transports, administration, citoyenneté, environnement, culture, éducation, santé,
- de la généralisation d'un accès aux jeux de données par type d'utilisation, autrement dit par profilage des acteurs concernés, ce qui en conditionne à la fois la visualisation et la catégorie d'utilisateur comme par exemple les plus fréquentes qui permettent de matérialiser les données sous la forme de tableau ou de carte ou d'analyses (abaques par exemple) mais aussi d'exporter directement ou d'accéder à l'API (*Application Programming Interface*).

Ces quelques éléments structurants témoignent des acquis de la première période de l'Open Data métropolitain et le stabilise, même si les réutilisateurs professionnels qui seront étudiés dans la suite de cet ouvrage soulignent toujours le manque cruel d'ho-

mogénéité et de pérennité des jeux de données. Mais il faut cependant les compléter par la réflexion sur les licences conduite par plusieurs métropoles.

Les licences sont l'élément clef du modèle économique mais elles témoignent aussi du degré d'appropriation des directives par l'acteur territorial porteur. Ici le cas de *Data Grand Lyon* est singulier et leader car il propose de revenir sur les licences validées par Etalab. En effet, dans un premier temps, suite à la prolifération anarchique des licences, la Loi pour une République numérique de 2016 a listé [18] les licences « *qui peuvent être utilisées par les administrations pour la réutilisation à titre gratuit de leurs informations publiques, qu'il s'agisse de données ou de code source d'un logiciel* » (article D.323-2-1 du code des relations entre le public et l'administration). Fort de son expérience et dans une démarche d'innovation, *Data Grand Lyon* a d'abord développé une licence complémentaire, différente et soumise à redevance. Une telle proposition montre à la fois l'évolution des portails et la réelle mise en place d'un écosystème qui cherche son ou ses modèles économiques dans une relation de concrétisation d'une idéologie dans les réalisations métropolitaines. Cette licence, dite « licence associée » a d'abord visé :

> *à favoriser le développement de nouveaux services et entreprises émergentes tout en garantissant un écosystème concurrentiel équitable (« fair business »). Pour ce faire, elle introduit une possibilité de redevance pour les acteurs offrant des services dépassant les seuils de monopole reconnus par la Commission européenne (40%). Cette licence n'intègre aucun frais fixes afin d'éviter toute barrière à l'entrée (pénalisante pour les petits acteurs) et se déclenche de manière exceptionnelle et exponentielle en fonction d'un seuil élevé de volume d'activité du produit ou du service commercialisé issu des données, traduisant la position que l'entreprise occupe sur son marché pour ledit service exploitant les données ouvertes. La redevance sera ainsi nulle sauf en cas de position tendant à devenir monopolistique. Par ailleurs, l'exonération est appliquée systématiquement pour les PME innovantes, le monde académique et un certain nombre d'usages ou réutilisations. La Métropole de Lyon agit au travers de cette redevance pour empêcher tout abus de position dominante et se positionne comme garante de l'absence de monopole, favorisant l'innovation.* [19]

Cette licence partait donc du principe qu'un modèle économique de ce type, redevance sélective, pouvait être appliqué à l'exploitation de certaines données métropolitaines, même si, dans les faits, à ce jour, « la licence "associée" n'a généré aucune redevance car aucun des réutilisateurs ayant signé cette licence (soit plus d'une centaine en continu chaque année) ne dépasse le seuil de « parts de marché » (volume d'activité) conduisant à une redevance effective ». Mais la Métropole de Lyon a encore franchi un pas supplémentaire en soumettant à homologation d'Etalab en juin 2019 sa « licence LRDIG ».

18. https ://www.data.gouv.fr/fr/licences, repéré le 4 mai 2019
19. https ://data.grandlyon.com/connaitre-nos-licences, repéré le 4 mai 2019

> *Cette licence de réutilisation des données d'intérêt général*[20] *(LRDIG) conserve en premier lieu le principe essentiel pour la Métropole de Lyon de garantir pour certaines données, une réutilisation qui ne va pas à l'encontre de l'intérêt général et des politiques publiques mises en œuvre sur le territoire, et qui prend en compte les évolutions légales (notamment la Loi pour une République numérique votée en 2016 et la Loi d'Orientation des Mobilités promulguée en décembre 2019), les remarques formulées par les services juridiques de l'État (échange entre les experts juridiques numériques métropolitains et leurs homologues), et les limites observées par les réutilisateurs, entendues par la Métropole de Lyon (notamment les questions relatives à un droit d'audit sur les systèmes d'information des réutilisateurs et la durée jugée trop courte de la licence nécessitant de la renouveler trop fréquemment, entre autres). La « LRDIG » a été votée le 30 septembre 2019 en même temps qu'un cadre élargi de licences comme prévu au code des relations du public avec les administrations*

Bien sûr, cette nouvelle licence LRDIG sert de pont avec les stratégies de transformation vers une ville intelligente car elle souligne la volonté ferme de structurer cet écosystème par les données et autour des données à la fois en direction de tous les réutilisateurs des plus modestes à ceux qui portent l'innovation et ses services à haute valeur-ajoutée et complète l'ouverture par le partage des données. S'il s'agit là d'une nouvelle phase du développement économique du territoire puisque les portails Open Data métropolitains constituent une action forte en termes de marketing territorial. La reconnaissance de cette licence unique en France à ce jour, témoigne enfin de la puissance de l'acteur métropolitain à structurer son propre écosystème, mais également à soutenir la concrétisation d'une idéologie en actions, décisions et principes communs de fonctionnement et donc peut-être même jusqu'à en porter la généralisation ou l'agrément au niveau national[21].

4 La métropole comme acteur fédérateur

Ainsi la métropole reste-t-elle l'acteur clef fédérateur et premier de l'ouverture des données à l'échelle du territoire, en lien étroit cependant avec deux autres catégories d'acteurs indispensables puisque ce sont ses cibles, et c'est pourquoi nos travaux les ont privilégiés : les réutilisateurs, programmeurs ou développeurs de métier, véritables artisans de la valorisation des données et les citoyens, théoriquement destinataires des services à venir.

Désignés comme développeurs sur les portails étrangers, on les appelle plutôt réutilisateurs dans le cadre hexagonal ; mais les deux dénominations correspondent bien à ces professionnels de la donnée qui développent des offres de traitement des données sur le principe d'un courtage informationnel automatique et programmé (applications) ou sur mesure (services). En aval de la métropole, qui leur ouvre souvent un accès privilégié dès la page d'accueil, ils assurent une fonction, sélective, de médiation des

20. https ://www.grandlyon.com/delibs/pdf/Conseil/2019/09/30/DELIBERATION/2019-3724.pdf, repéré le 31 octobre 2019

21. Cette partie a été nourrie des échanges passionnants avec Madame Vernus-Prost que je remercie ici pour sa disponibilité et son intérêt pour ce travail.

données et transforment la data en information. C'est bien ce processus de transformation qui est au cœur de l'Open Data puisqu'il porte la valorisation des données et surtout en rend possible, ou facilite en tout cas, son appropriation et son utilisation par les citoyens. Dans le livrable déjà mentionné (Paquienséguy et Dymytrova : 2017), nous remarquions justement à l'analyse de cinq des portails étudiés (Londres, Berlin, Séoul, Chicago, New-York) qu'en isolant l'accès développeurs, les métropoles valorisent encore plus le citoyen, destinataire premier de tous les autres rubriques proposées. En effet, les choix des menus destinés uniquement aux individus se trouvent alors débarrassés de toute considération technique de format, de modalités d'accès ou de réemploi, et leurs usages s'en trouvent simplifiés. Mais ce critère d'un profilage des destinataires du portail dès sa page d'accueil est insuffisant à considérer la place des développeurs, car d'autres villes étudiées (Lyon, Berlin, Montréal) les prennent en compte différemment, en amont, au travers des licences, non marchandes ou commerciales des données, comme nous l'avons dit, elles posent alors le contexte économique et juridique du réemploi des données publiques ouvertes.

Autrement dit, trois éléments issus du positionnement de la métropole pour ouvrir ses données ou plutôt les données du territoire, structurent la place et définissent les opportunités offertes aux développeurs et data journalistes. Premièrement, la place qui leur est donnée, ou réservée, dans l'architecture même de la plateforme et qui vérifie, ou pas, un accès prioritaire aux jeux de données ; ici le portail berlinois serait l'exemple caractéristique. Deuxièmement, les licences qui accompagnent, ou pas, leur travail et leurs succès en proposant un modèle économique de valorisation des données sur le principe de la redevance, ici donc les données ouvertes affirment leur potentiel à générer de la valeur à toute échelle, et pas seulement à une échelle de Big Data ; plusieurs développeurs ont d'ailleurs confié en interview n'utiliser que quelques jeux de données des portails métropolitains ; le cas de *Data Grand Lyon* est bien entendu le plus représentatif ici. Troisièmement, l'affichage, ou pas, des applications développées directement sur le portail métropolitain afin de faire le lien direct entre les données et leur exploitation sous forme de prestation de services à tous ; et c'est sans doute ici que les portails français sont le plus en retrait au regard d'autres portails internationaux, celui de Séoul par exemple qui a une vraie stratégie de présentation et de promotion des applications et des services conçus à partir de ses données.

Finalement, les développeurs constituent une catégorie d'acteurs également hétéroclite puisqu'il s'agit ici à la fois de véritables professionnels, de geeks aux idéaux de libre, de jeunes innovateurs à la recherche de l'application à succès, dont les pratiques professionnelles et univers sont variés, en termes de langage de codage utilisé, en termes de compétences, en termes de routine quant aux sources exploitées pour extraire des jeux de données. En effet, cette catégorie finalement floue des « développeurs » intervient dans deux registres bien différents : certains développeurs exploitent les données métropolitaines pour étendre leurs propres activités alors que d'autres proposent, à partir des mêmes données ouvertes, des applications gratuites au service de tout un chacun. Sans parler de choc entre les deux registres il faut bien remarquer une grande tension sur fond de désaccord ou de complémentarité qui témoigne du caractère biface des portails métropolitains Open Data qui par une même réalisation cherchent aujourd'hui à tenir à la fois de l'idéologie initiale de l'ouverture des données,voire du

code, en favorisant accessibilité, gratuité, et réutilisation des données métropolitaines pour tous, et à la fois de ressources à monétiser, à valoriser par la publication, l'exploitation et la transformation des données par des professionnels de la data, divers et variés par ailleurs.

Si le cas de *Data Grand Lyon* reste encore isolé, voire unique avec la licence associée, pour rappel assortie d'une redevance en cas de succès économique, c'est bien cet exemple qui témoigne de cette dualité d'objectifs et que le portail génère une monétarisation des données qui constitue un apport, certes insuffisant, mais non négligeable au regard des coûts fixes liés à l'Open Data métropolitain. D'ailleurs, il faut souligner que d'autres portails comme ceux de Berlin et Montréal réfléchissent à des solutions de compensations également. Par ailleurs, ces applications sous licence associée, dynamisent l'offre Open Data et l'inscrivent encore plus soit dans une démarche *Smart City*, soit dans la promotion du territoire.

Cependant, la thématique de l'ouverture des données impose également d'enrôler une autre catégorie d'acteurs, centrale et atypique que nous appellerons celle des citoyens. Difficile en effet de la nommer et de l'identifier puisqu'il s'agit des utilisateurs, bénéficiaires et destinataires des données ouvertes, et finalement le terme citoyen porté par le processus d'ouverture lui-même est sans doute le plus adapté, même s'il ne rend compte que de cette idéologie et non de la réalité des cibles définies par les portails métropolitains. Comment oublier en effet, que l'Open Data porte le flambeau d'une idéologie d'ouverture, de participation, de transparence. Les données sont ouvertes pour eux. L'accès de droit à l'information et aux données publiques, déjà voté en France depuis 1978 [22], et plus largement promu par le Programme d'Action Gouvernementale pour la Société d'Information (Pagsi) lancé par Lionel Jospin en 1997 ont rendu les administrations, et les métropoles avec elles, responsables de l'ouverture de leurs données avec comme destinataire premier et unique en principe le citoyen. Mais par ailleurs, la catégorie des citoyens est atypique car elle représente le seul type d'acteur à ne pas prendre part directement aux activités économiques générées ou nourries par l'Open Data puisqu'il ne vérifie pas d'intéressement marchand ; et à militer, collectivement *via* des associations de citoyens telles que Libertic, la Fing ou Regards Citoyens par exemple, pour une ouverture des données permettant pour le moins, facilitant pour le plus, leur utilisation par le simple citoyen en vue d'une appropriation dite citoyenne des données. Deux éléments cristallisent la place, ou l'absence, du citoyen dans ces portails : les thématiques retenues pour classer et présenter les jeux de données, et les modalités choisies pour les lire. D'ailleurs, certains portails opèrent de façon flagrante une médiation au profit du citoyen afin de lui faciliter la tâche en développant à la fois des outils de visualisation des données et un premier traitement produisant une valeur ajoutée directement consultable. Ainsi, Londres, Chicago et Séoul offrent-elles une visualisation sous forme de graphiques et figures, lisibles par tout un chacun, sans compétence technique aucune même si bien entendu,

22. 20 Loi n° 78-753 du 17 juillet 1978 reconnaît à toute personne le droit d'obtenir communication des documents détenus dans le cadre de sa mission de service public par une administration, quels que soient leur forme ou leur support. Ce droit s'exerce à l'égard de toutes les personnes publiques (l'État, les collectivités territoriales et leurs établissements publics) ainsi qu'à l'égard des organismes privés chargés d'une mission de service public.

toute figure qui schématise et modélise les données et la réalité sociale et économique qu'elle représente doit être interprétée.

De même, l'intitulé et le nombre de thématiques montrent le degré de prise en compte des intérêts directs du citadin. Si la thématique la plus présente est celle de l'environnement, la plupart des jeux concernent le citoyen et l'informent sur la qualité de l'eau, de l'air, sur les avis sanitaires quant aux lieux de restauration, sur l'état du déneigement des rues, les places libres dans les parkings souterrains, mais *via* des modalités qui ne lui sont pas adaptées : par exemple, Montréal propose des données sur la qualité bactériologique et physicochimique des ruisseaux et plans d'eau intérieurs de l'agglomération qu'un citoyen ne peut s'approprier sauf à avoir des compétences particulières en chimie ; Chicago informe sur les inspections alimentaires des lieux de restauration de la ville sur la base de nombreux graphiques (simples abaques ou courbes) qui traitent les données de façon globale et annuelle sans permettre d'identifier les restaurants accrédités par la *Food Administration* pour autant, le Grand Lyon propose des places de parking qui restent libres 30 secondes en moyenne. Grâce à de simples courbes, Londres tient le compte du ratio du poids des ordures ménagères par habitant en comparant celui de la ville à celui du pays, certes le citoyen en sera informé mais sans pouvoir en tirer aucune conclusion pour autant. Dans la très grande majorité des thématiques, faire faire sens aux jeux de données afin d'en tirer une information utile n'est pas chose simple pour le citoyen car la structuration des données d'une part et leur mise en forme de l'autre ne lui sont pas destinées.

Le contraste est flagrant avec les jeux portant une volonté manifeste de lui fournir des informations exploitables et non des données brutes. Lyon par exemple liste les sentiers nature de l'agglomération, en montre les tracés dès le premier clic et fournit l'itinéraire au format PDF. En deux clics, le marcheur se décide sur carte et plan familiers, semblables aux topos habituels. *Il en va de même* pour les silos à verre dont les emplacements sont signalés sur un plan. Cependant, un niveau supplémentaire de traitement des données vient des applications que proposent Amsterdam, Berlin, Séoul et Montréal. Elles fournissent le dispositif d'intelligence d'accès aux données et en assurent une médiation destinée aux citoyens ou résidents par le biais d'une prestation de services. Séoul par exemple, propose de véritables applications [23], également disponibles sur l'*App Store*, qui offrent des services liés à la fréquentation des transports publics, depuis éviter les stations de métro les plus bondées [24], aux parkings, aux stations de gonflage et à la location de vélos municipaux. Quant à l'application *Neo Séoul*, elle répertorie les lieux et événements culturels.

L'agencement de l'alliance stratégique entre acteurs montre la chaîne partenariale qui se met en place et l'écosystème qui se construit. Le portail de Séoul, comme bien d'autres, permet à des développeurs de se saisir des données métropolitaines afin de créer des applications qui constituent un degré supplémentaire de médiation envers l'utilisateur final (citoyen, résident, touriste, etc.). Ces applications valorisent à la fois la ville, ses performances, le développeur et les données. Plus largement distribuées et

23. Disponibles dès la page d'accueil du portail http ://data.seoul.go.kr/, repéré le 20 novembre 2019

24. Le métro de Séoul est particulièrement surchargé et voit passer 2,8 millions de voyageurs par jour et par ligne ce qui rend cette application véritablement stratégique à la fois pour les voyageurs et *SéoulMétro*. (Communiqué de presse de l'agence YonHap, 21 juillet 2015)

diffusées, elles sont doublement accessibles par le portail métropolitain et sur l'App Store qui assure lui aussi un (autre) niveau de médiation et de propagation. Ainsi, l'alliance stratégique des partenaires fonctionne puisque chacun en retire un bénéfice adapté : le citoyen accède à une information utile et opérationnelle qui éclaire ses choix, le développeur accède à la notoriété *via* le nombre de téléchargements de son application et la ville s'en trouve à la fois mieux connue et plus accessible (encombrement du métro par exemple).

Cet équilibre fonctionne grâce à la place prise ou négociée par les développeurs dans le processus d'ouverture des données métropolitaines.

5 En route vers l'industrialisation des données ouvertes ?

La logique d'*empowerment* déployée par la métropole se matérialise clairement ici. En effet, au-delà de ces simples avantages partagés, il semble que les métropoles ayant d'ores et déjà accompli un processus d'intégration et de capacitation des acteurs producteurs d'applications ou réutilisateurs des données de façon nette et structurelle dans les portails Open Data métropolitains s'engagent dans un processus de structuration organisationnelle.

Ainsi, la concrétisation des idéaux de l'ouverture des données produira-t-elle des résultats à la fois différents quant aux productions, applications, utilisations en présence, et des processus assez similaires de fédération d'acteurs et de compétences. Dans la première décade, à l'évidence, c'est la diversité qui règne parmi les portails métropolitains qui, tous, cherchent, essaient, expérimentent, leurs multiples versions qui se suivent, leurs dénominations qui évoluent en témoignent. Mais ensuite, dans l'actuelle deuxième décade, les traits communs de la structuration organisationnelle de l'Open Data, de ses outils et de ses acteurs apparaissent même s'ils donnent lieu à des actions spécifiques comme nous venons d'essayer de le montrer.

Les trois caractéristiques de l'industrialisation (Mœglin : 2016) se retrouvent présentes ici et offrent une grille de lecture intéressante pour saisir le niveau de structuration de l'Open Data, autrement dit ses avancées vers une industrialisation, laquelle sous-entend un ou des modèles économiques en présence, y compris celui de la gratuité pour l'utilisateur avec la licence ouverte, des valeurs communes, et une certaine pérennité. En ce qui concerne l'ouverture des données, le premier trait, l'idéologisation, qui fédère les acteurs à l'œuvre, est déjà marqué, le début de ce chapitre a cherché à le montrer. En effet, comme le résumait simplement Jean-François Tétu il s'agit par l'idéologisation d'« *imposer l'idée que ce processus arrange tout le monde* » (Tétu : 2016, 568) sur la base d'arrangements différents, spécifiques qui dépendent des acteurs et des volontés en présence ; ce qui explique d'une part que les portails varient d'une métropole à l'autre tant au niveau des choix et des formats proposés qu'au niveau de la vision de l'Open Data qu'ils vérifient. Le deuxième trait est celui de la technologisation qui paraît être l'élément central et la préoccupation première aujourd'hui, même si « *l'essentiel est bien l'interdépendance entre ces trois traits* » (*ibid.*, 569) ce qui nous autorise à en inverser l'ordre proposé afin de mieux s'adapter à la réalité observée. La technologisation façonne les pratiques, ici professionnelles principalement, et les façons de faire des acteurs sur la base de deux marqueurs (*ibid.*, 568) qui entraîneront

l'industrialisation : 1/ un rationalisme technique et 2/ une organisation productive. Dans le cas de données ouvertes, le premier marqueur se trouve fortement facilité par bien des éléments et s'exprime forcément dans des formes différentes que dans le cas de l'industrialisation de l'éducation (Mœglin : 2016) car la nature même de l'activité y pousse, ou du moins facilite ce rationalisme (routine informatique, formatage, extractions automatiques, catalogue de données, thématiques des jeux de données, etc.). Cependant, certaines catégories de professionnels de la data (Dymytrova, Larroche et Paquienséguy : 2018) travaillent encore manuellement, de façon artisanale et experte pour nettoyer ou qualifier des données afin de les formater pour leurs propres API ; étape indispensable à cause de leur hétérogénéité (sources, formats, structuration, etc.). Le travail conduit dans l'ANR *OpenSensingCity* cherchait d'ailleurs à contribuer à cette rationalisation en proposant une couche logicielle d'uniformisation des formats et qualités de données pour en faciliter l'exploitation [25]. Quant à l'organisation productive, il faut d'abord remarquer que la création récente (à partir de 2016 environ) dans les services des métropoles du poste de *Chief Data Officer* formalise leur volonté de soutenir et d'accompagner cette mise en place organisationnelle, principalement avec les acteurs dominants sur leur territoire (industriels, universités, collectifs, mairies de la métropole, régies municipales de transports, etc.). Mais l'organisation productive passe aussi par deux éléments qui ont déjà été évoqués : les licences et le profilage du portail. Finalement, la technologisation de l'Open Data correspond bien à un processus de « *mise en technologie d'un projet politique* » (Badouard : 2014), celui de l'ouverture des données.

La force de la technologisation exprimée dans les portails métropolitains conduit à marquer le troisième trait, celui de la rationalisation. En effet, pour l'Open Data et ses portails, elle passe d'abord par le profilage (citoyens, geeks, développeurs) qui entraîne une stigmatisation des compétences et une séparation des tâches.

Si pour les professionnels de la data l'affaire semble évidente, les citoyens n'en sont pas exclus car leurs actions relèvent d'un *digital labor* à la fois dans la production du service (self-service, coproduction, paramétrage, etc.) et dans la production des données (géolocalisation, cookies, adhésion, abonnements).

De plus, comme pour l'industrialisation de l'éducation (Mœglin : 2016), ici la rationalisation est principalement portée par les acteurs publics, représentant la fonction publique territoriale, et en amont par les institutions étatiques formant l'écosystème de l'Open Data à la française (Mabi : 2015) autour des acteurs dominants tels Etalab ou le portail data.gouv.fr. La rationalisation de l'ouverture des données est donc un processus actif et efficace car, au-delà des acteurs publics territoriaux ayant finalement fonction de maître d'œuvre, les différentes catégories de professionnels de la data qui gravitent autour des portails métropolitains, développeurs, *data scientists* et data journalistes, ont déjà atteint un niveau de rationalisation assez structuré et pérenne pour produire une chaîne de traitement de données générant une création de valeur ajoutée.

25. Couche logicielle développée par l'EMSE présentée ici https ://github.com/OpenSensingCity, repéré le 10 juillet 2019.

Les bases sont alors suffisamment stables pour donner naissance à un ou des modèles économiques de l'Open Data.

5.1 Le courtage informationnel comme modèle économique de référence ?

Finalement, à l'échelle internationale le constat se dresse que : « *construire un portail Open Data est coûteux, et [que] les retours sur investissements peuvent être difficiles à mesurer*[26] » (Murray : 2017). Ainsi, l'ampleur du portail et la visibilité qui lui sont données relèvent surtout des stratégies discursives et organisationnelles des acteurs publics, d'Etalab aux métropoles et à leurs professionnels de la data. Le premier livrable (Paquienséguy et Dymytrova : 2016) [27] de notre recherche et sa synthèse (Paquienséguy et Dymytrova, 2017), retraçant l'analyse de vingt-quatre portails métropolitains Open Data, soulignent à quel point la réutilisation des données ouvertes reste l'objectif premier et officiel des métropoles, lequel conditionne la viabilité du portail. De façon transverse, se retrouve la matérialisation de la réutilisation des données ouvertes dans une application ou une fonctionnalité à destination du citoyen, du résidant ou du touriste car pour

> *assurer la participation des citoyens aux portails de données ouvertes, les responsables municipaux et les partenaires-tiers conçoivent des interfaces et des applications axées sur applications Web et mobiles préférées, conçues pour aider les citoyens à naviguer dans les données ouvertes*[28] (Murray : 2017).

Autrement dit, précaire et difficile, l'équilibre économique des portails Open Data doit se trouver entre les idéaux de transparence et ceux de création de valeur ajoutée, lesquels se matérialisent à la fois par les choix de la gouvernance locale d'ouvrir certains jeux de données ou pas, et par ceux des applications proposées, retenues ou soutenues par l'acteur pivot, *via* des hackathons ou des expérimentations par exemple. Cependant, les activités de terrain que nous avons conduites – enquêtes, workshops, expérimentations – et tout particulièrement l'ensemble des entretiens menés (n=59) montrent que les données ouvertes sont rarement travaillées seules, ou prises en compte comme telles. La façon dont les professionnels de la data (développeurs, réutilisateurs, data journalistes, *data scientists*, géomaticiens) en parlent et s'en saisissent en témoigne. En effet, notre troisième livrable de l'ANR (Dymytrova, Larroche et Paquienséguy : 2018 [29]) qui analyse leurs pratiques professionnelles en tant que réutilisateurs des données ouvertes fait ressortir les limites de l'exploitation et de la réutilisation des données ouvertes. Ces limites viennent à la fois des données

26. "*Building an Open Data portal is expensive, and returns on this investment can be difficult to measure"* traduction de l'auteur, https ://carto.com/blog/forty-brilliant-open-data-projects-preparing-smart-cities-2018/#navigating, repéré le 6 décembre 2019

27. Disponible ici : https ://hal.archives-ouvertes.fr/hal-01449348, repéré le 19 mai 2019

28. "To ensure citizen engagement with Open Data portals city officials and third-party partners are designing constituent-centric interfaces and applications. Here are our favorite Web and mobile applications designed to help *citizens navigate Open Data*", traduction de l'auteur, https ://carto.com/blog/forty-brilliant-open-data-projects-preparing-smart-cities-2018/#navigating, repéré le 10 décembre 2019

29. Disponible ici : https ://www.researchgate.net/publication/324976707*Livrable*3, repéré le 4 octobre 2019

ouvertes elles-mêmes et de la chaîne de production de la valeur mise en œuvre par les différentes catégories de professionnels. Autrement dit, les industries de la data, comme celles des contenus ou de la culture se trouvent avant tout régies et structurées par leur(s) propre(s) modèle(s) économique(s) de création de valeur. L'écosystème est complexe et pour partie instable, on le voit bien. D'une part les acteurs sont nombreux, certains d'envergure uniquement régionale, comme ceux de la gouvernance locale par exemple, mais d'autres d'ampleur internationale, comme ceux de la fourniture d'accès ou du traitement des données (Cisco ou Socrate par exemple) ; d'autre part, les données proviennent de différentes sources et répondent donc à des exigences et variables différentes en termes de mise à jour, de format, de licence ou de *privacy* ; sans oublier que les objectifs des données métropolitaines sont toujours duaux puisqu'ils s'inscrivent dans le même temps dans le processus d'ouverture des données et de transition numérique de la ville vers une *Smart City* à chaque fois repensée.

C'est pourquoi la réflexion sur le modèle économique de l'Open Data commence par le modèle de référence, c'est-à-dire celui du courtage informationnel (Mœglin : 2004), mais pour mieux rendre compte de la complexité que nous venons d'exposer, il doit être élargi aux, ou intégrer, les règles de la computation (Francony : 2017) qui sont le moyen de transformer la data en information [30], autrement dit, de créer sa valeur ajoutée.

Dès 2016, dans une première réflexion sur l'exploitation des données ouvertes selon le modèle du courtage informationnel (Mœglin : 2004, 225-227 ; 2005), nous avancions déjà l'adéquation des quatre éléments prégnants du modèle avec l'activité des professionnels de la data :

l'entremise [...] assure la fonction centrale ; la valorisation de l'intermédiaire se fait sur des modes complémentaires qui vont de la vente de mots-clés au paiement à l'acte ; le courtage lui-même est assuré par des logiciels oeuvrant à l'indexation comme à la recherche des données et au profilage perpétuel du consommateur [...] ; l'intermédiaire est valorisé pour lui-même, au-delà de la fonction centrale qu'il assure. (Paquienséguy : 2016, 354)

En effet, inscrire l'analyse des portails Open Data dans la logique du courtage informationnel conduit à un autre regard sur la place et les enjeux des collectivités territoriales à l'œuvre comme sur la place et le rôle des données ouvertes. Si un tel modèle convient à l'industrie de la donnée c'est d'abord pour plusieurs raisons génériques qu'il faut rappeler ici par souci de clarté et de précision. D'abord, « *il n'y a pas de concordance parfaite* » (Thuillas : 2017, 41) entre une filière ou industrie et un modèle, tels que ces derniers ont été définis pour les industries culturelles et médiatiques à partir des années quatre-vingt. Ensuite, les « *modèles ainsi définis connaissent de nombreuses hybridations ou remises en cause* » (*ibid.*) car sans déroger à leurs caractéristiques fondamentales, ils évoluent, s'adaptent aux mutations des filières afin de diversifier les modes de création de valeur. Enfin, les postes clés de la filière pour le bon fonctionnement de son ou ses modèles économiques sont parfois redistribués à

30. La transformation de la date en information sur la base d'un processus computationnel est présentée dans le chapitre conclusif de cet ouvrage De l'Open Data à l'information, quelques points d'appui depuis les SHS

plusieurs reprises avant qu'une forme de stabilité n'engendre une version suffisamment résistante du modèle pour qu'elle puisse en être rapprochée, voire associée comme ici. Cependant, la faiblesse majeure de la démonstration vient de la notion de filière : certes les données constituent véritablement une industrie particulière *via* l'oligopole, des géants de l'Internet et des data compagnies, voire même une forme d'industrie (manufacturière) particulière, l'Industrie 4.0 qui s'appuie centralement sur les données et l'analytique (Rapport McKinsey : 2015). Si plusieurs marqueurs témoignent de l'industrie de la donnée, la démonstration reste à faire que, sur le temps long, l'Open Data en constitue une filière.

5.2 L'Open Data comme filière de l'industrie de la donnée ?

La question se pose vraiment car les données ouvertes ne représentent qu'une infime part des données et de l'économie des données, tant du côté de leur production que de celui de leur exploitation, au regard des autres catégories de données (données personnelles, données confidentielles, données privées des entreprises, données vendues, données inutilisables, données cryptées, etc.). En fait, seuls les portails métropolitains les valorisent, sans cependant chercher forcément à les regrouper ou à les harmoniser. Nous l'avons déjà souligné, c'est l'hétérogénéité des objectifs et des méthodes d'ouverture qui a initialement prévalu puisque l'objectif premier n'était pas de générer de la valeur, mais la transparence et la mise à disposition, bref, l'ouverture ! La fragmentation territoriale des données ouvertes et les spécificités locales des portails métropolitains freinent, bloquent le plus souvent, leur exploitation à l'échelle industrielle, d'où la présence d'autres types d'acteurs, plus souples, moins systématiques qui utilisent des données de niche (quelques jeux seulement) pour un développement ciblé et limité, à un territoire dans la plupart des cas, comme les start-up, coopératives, incubateurs, *factory, lab* ou autre type de (petites) structures qui peuvent s'analyser comme les prémices de la structuration d'une filière.

Comme pour la musique enregistrée, première industrie culturelle à se transformer rapidement et en profondeur au contact des nouvelles pratiques culturelles et des technologies numériques qui s'y rapportent, se retrouve donc pour une potentielle filière des données ouvertes « *une myriade de petites structures, plus souples et plus aptes à dénicher de nouveaux talents et produisant un grand nombre de contenus qui pour l'essentiel sont diffusés par les grands acteurs* » (Bouquillion : 2012, 2) Les talents seraient ici les applications ou les idées d'applications d'abord développés dans un contexte limité de données, de lieux et d'utilisateurs. Ainsi, si ces acteurs ne se situent pas en amont des acteurs industriels de poids, ni même tout au long du processus industriel comme c'est souvent le cas pour une filière, ils ne se situent pas non plus complètement à l'aval. Ils occupent une phase intermédiaire de l'aval, porteuse de l'idée et du concept qui ne peut passer à l'échelle que sous couvert d'un rachat industriel, d'un passage à l'échelle (de start-up à scale-up) ou d'un changement de statut. Cependant, vu leur nombre, le vivier et le levier qu'ils constituent, leur place et leur rôle ne sont pas négligeables dans l'écosystème général de l'Open Data. Les métropoles par exemple les sollicitent et encouragent principalement par le biais de laboratoires urbains et lieux d'innovation et d'expérimentations urbaines. Ces *Urban Labs* constituent des acteurs tiers nouveaux qui se multiplient, dans les métropoles

principalement, jusqu'à faire modèle[31]. Dans notre étude nous avons par exemple remarqué certains de ces acteurs comme le *Tubà* lyonnais, qui se définit comme un « *lieu d'innovation et d'expérimentation pour la ville de demain, est porté par l'association Lyon Urban Data. Il favorise l'innovation, l'incubation et le développement de services urbains s'appuyant sur les données numériques privées et publiques* »[32] ; la *Turbine* grenobloise qui se présente comme un « *infolab pour outiller l'innovation socio-économique et les données*[33] » ; ou encore le *Nantes City Lab* qui « *met à votre disposition un panel de sites d'expérimentation, d'équipements, de données et d'ingénierie pour tester et accélérer votre concept grandeur nature*[34] ».

Autrement dit, les acteurs de la filière de l'Open Data commencent à être identifiables aujourd'hui. La métropole, ses services et directions, ses régies et partenaires en sont les plus visibles, tout comme les *Urban Labs* et autres structures intermédiaires entre la collectivité territoriale et les citoyens. Ces tiers lieux portés par les métropoles ou indépendants, car les deux registres existent, favorisent et soutiennent les démarches citoyennes. Complétés par les industriels disposés à investir ces questions, de force lorsqu'ils représentent un des acteurs clefs du secteur comme par exemple Decaux pour la mobilité urbaine, ou par choix lorsqu'ils disposent de donner à partager librement, ils forment les acteurs phares, leaders de l'Open Data du territoire et construisent un écosystème dont la spécificité révèle la fragilité ; d'autant plus que la création de valeur reste incertaine.

5.3 Un modèle économique de référence, ou plusieurs ?

Quel que soit le secteur d'activité auquel on l'applique, un *business model* désigne une organisation dont l'objectif est de générer de la valeur[35] de façon pérenne et de telle sorte que l'ensemble des acteurs pris dans cette organisation en bénéficient, dans des proportions variables et selon des modalités ajustées. Un *business model* identifie donc à la fois les coûts et les revenus d'une part et les acteurs qui les supportent ou génèrent de l'autre. La filière de l'Open Data, plus précisément celle de l'*Open Government Data*[36], a pour trait particulier de désigner *de facto* les acteurs devant supporter les coûts d'infrastructure et de collecte puisque ce sont les métropoles qui développent les portails Open Data et les collectivités territoriales de plus de 3500 habitants qui fournissent les données ouvertes. Mais c'est encore l'acteur métropolitain qui se doit de collecter et traiter les données pour une première harmonisation,

31. Raphaël Besson, « Urban Labs, un modèle sous haute tension », septembre 2018, *The Conversation*, https ://theconversation.com/urban-labs-un-modele-sous-haute-tension-103039, repéré le 2 juin 2019

32. http ://www.tuba-lyon.com/tuba-cest-quoi/, repéré le 6 juin 2019

33. https ://turbine.coop/infolab/, repéré le 6 juin 2019

34. https ://www.nantesmetropole.fr/pratique/numerique/nantes-citylab-et-si-nantes-devenait-votre-terrain-de-jeu-developpement-durable-urbanisme-92851.kjsp, repéré le 8 juin 2019

35. Alexander Osterwalter Yves Pigneur, *Business Model Generation*, repéré le 4 avril 2019, https://s3.amazonaws.com/academia.edu.documents/32253198/businessmodelgenerationpreview.pdf?response-content-disposition=inline\%3B\%20filename\%3DYoure_holding_a_handbook_for_visionaries.pdf\&X-Amz-Algorithm=AWS4-HMAC-SHA256\&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A\%2F20190606\%2Fus-east-1\%2Fs3\%2Faws4_request\&X-Amz-Date=20190606T141325Z\&X-Amz-Expires=3600\&X-Amz-SignedHeaders=host\&X-Amz-Signature=11130dc09e292a47fccfbbc534eeec2f6b32013b5d20c7077a18d599ba8cbc85

36. « *Open Government Data* » et « *Open Data* » en général, car la distinction est claire dans la littérature anglo-saxonne . En effet, beaucoup d'entreprises créent aussi leurs portails Open Data pour valoriser les données qu'elles publient gratuitement et inciter à leur réutilisation.

avant de les mettre à disposition. Par contre, il est évident que dès lors il en maîtrise la publication et les conditions d'ouverture, autrement dit l'éditorialisation comme le montre le dernier chapitre de l'ouvrage ; certaines métropoles saisissent là l'opportunité de nouer ou contracter des partenariats avec les acteurs majeurs de la mobilité urbaine comme Decaux ou Vinci. Si la diversité des situations et des ambitions que les portails Open Data métropolitains matérialisent, et que notre analyse de terrain a montrée, n'est plus à mettre en doute, les modèles économiques d'exploitation de l'Open Data semblent converger. Nous retenons deux courants majeurs pour les déterminer. En effet, même si les portails se renouvellent fréquemment (environ tous les deux ans) dans une recherche encore intense de la solution la plus adaptée, ils se rejoignent principalement d'une part sur les trois modèles économiques de l'*Open Data Institute* (ODI) initialement proposés par la fondation de Tim Berners-Lee [37] en 2013, et de l'autre sur les trois hypothèses instruites par l'ouvrage *Datanomics, les nouveaux business models des données,* de Simon Chignard et Louis-David Benyayer, paru deux ans plus tard de l'autre côté de l'Atlantique.

C'est dans son célèbre billet intitulé "*How to make a business case for Open Data*", que T. Berners-Lee suggère « *en première instance d'identifier les buts que l'organisation cherche à atteindre avec l'ouverture des données* » [38] ; avant de décliner les trois modèles de création de valeur à partir de l'ouverture des données. Pour rappel [39], le modèle *freemium* se base sur la fourniture d'un service ou d'un produit à valeur ajoutée, facturé au bénéficiaire ; en essence, il est le plus proche du modèle du courtage informationnel car il repose véritablement sur une prestation service spécifique qui ajoute de la valeur à l'information ou aux données fournies. Le modèle de financement croisé compte beaucoup plus de partenaires ou d'utilisateurs avec lesquels les services sont plus profilés, voire sur mesure en recourant à des donnés partagées ou plus utilisées ; il se réfère également au modèle du courtage informationnel mais avec l'idée d'une industrialisation sous-jacente qui permettra un passage clef à l'échelle, des clients comme des datas. Enfin, le modèle également réticulaire et collaboratif regroupe plusieurs organisations afin de réduire les coûts fixes et d'élargir le public potentiel des services et produits à valeur ajoutée. Ce modèle entre complètement en résonance avec celui développé par les métropoles qui se sont associées aux grands acteurs de leurs territoires (souvent dans le secteur des transports ou du tourisme) afin d'une part de fédérer toutes les données locales et de l'autre d'offrir un service continu et si possible en temps réel.

S. Chignard et L.-D. Benyayer formulent quant à eux trois hypothèses pour comprendre pourquoi les modèles économiques de la donnée ouverte achoppent alors même que l'économie (mondiale) de la donnée s'amplifie rapidement à partir des principes de l'économie biface (Rochet, Tirole : 2003) qui irradie les plateformes d'intermédiation

37. https ://theodi.org/article/how-to-make-a-business-case-for-open-data/, repéré le 17 juin 2019

38. "*The first step in working on a business case for Open Data is to identify your organization's goals*"

39. « *Freemium : you provide an "added value" data product or service, for which you charge ; cross subsidy : you reach more customers, or provide enhanced services to existing customers, through wider sharing and use of your data ; network effects : by collaborating with other organizations, you reduce your costs in maintaining data which you use in your work or extend the possible audience for your products and services* ». https ://theodi.org/article/how-to-make-a-business-case-for-open-data/, repéré le 17 juin 2019

et les applications gratuites ; à partir de la curation systématique et algorithmique qui soutient les réseaux sociaux et les plateformes collaboratives.

Ils montrent [40] dans un premier temps que le temps de l'Open Data n'est pas encore là, et « qu'il est trop tôt », ce que nos travaux confirment en mettant l'accent sur l'effort de structuration en cours dans les services métropolitains, en soulignant la diversité des cas étudiés que nous regroupons sous trois types (Paquienséguy et Dymytrova : 2019) afin d'identifier les spécificités récurrentes qui sont sans doute un des éléments de structuration lorsqu'elles deviennent transversales, en remarquant progressivement qu'une panoplie (Labelle : 2007) se déploie dans les métropoles afin de soutenir l'ouverture, malgré la diversité des objectifs et des ambitions.

Dans un deuxième temps, et nos travaux les rejoignent ici sans doute encore plus, c'est le caractère diffus et complexe des données ouvertes qui constituerait un verrou fort. Effectivement, comme la grande majorité des réutilisateurs interrogés (Paquienséguy et Dymytrova : 2018) l'ont exprimé, les données ouvertes ne sont qu'une des sources de jeux de données utilisés qui regroupent les données crowdsourcées, les données payantes, des données d'entreprises... Autrement dit, les données ouvertes se trouvent englobées, voire noyées, dans les autres types de données ; à cette échelle du Big Data elles restent mineures et complémentaires pour les réutilisateurs du monde économique, même si elles sont centrales pour ceux du monde académique. Les isoler, pour leur attribuer un rôle ou un poids dans un modèle économique reste difficile, et sans doute vain, car leur découvrabilité [41] représente toujours un frein majeur aujourd'hui malgré les efforts d'Etalab.

En dernier lieu, nos travaux s'intéressent à un type d'acteur déterminant que nous avons travaillé à comprendre et à cerner : les réutilisateurs. Les start-up tant attendues pour développer des applications innovantes à partir des données publiques ouvertes et ainsi proposer une chaîne de création de valeur viable sont à la fois peu nombreuses et temporaires, le succès conduisant au rachat, dans la logique des licornes, l'échec à la disparition. Pour les réutilisateurs citoyens, c'est encore trop tôt comme le stigmatise la première hypothèse, les collectifs existent, accueillis par les tiers lieux. Notons d'ailleurs que ces lieux de partage, de coopération et d'innovation citoyenne, bénéficient dans la plupart des cas nationaux soit du soutien de la métropole, soit de celui de l'Union Européenne, *via* les fonds Feder tout particulièrement comme Hôtel71 porté par Arty-Farty à Lyon, ou Bordeaux Data Sciences dans la métropole aquitaine ou encore Eclozr à Rennes.

Une fois ces trois hypothèses posées, S. Chignard et L.-D. Benyayer ne distinguent finalement pas les Open Data des autres données et écrivent que « *les caractéristiques de la valeur de l'open data sont celles des données en général : la valeur est dans la réutilisation, elle est future et co-construite* » (Chignard et Benyayer :2015). Autrement dit, les données ouvertes ne porteraient pas de modèle économique spécifique mais s'intégreraient aux modèles de l'économie des données et de ceux de l'économie

40. https ://donneesouvertes.info/2014/07/28/datanomics-open-data-show-me-the-money, repéré le 28 juin 2019

41. https ://donneesouvertes.info/category/articles/economie-et-strategie-open-data/, repéré le 28 juin 2019

des plateformes soit sur un principe d'exploitation (la donnée comme ressource, le cadastre, les parkings, les silos à verre, etc.), soit sur un principe de production (la donnée industrialisée, issue de capteurs, de métriques etc.), soit sur un principe idéologique (la donnée ouverte, offerte au citoyen par les portails Open Data en vue de son partage et de sa réutilisation). Le décalage entre les discours et injonctions sur l'ouverture et la part des données ouvertes dans l'économie des données montre bien qu'il s'agit plus à ce jour encore d'une idéologie que d'une industrialisation, même dans les métropoles qui les destinent avant tout aux citoyens et ses collectifs, comme Rennes, Montpellier, Nantes ou Lille par exemple. Des analyses des spécialistes engagés de l'Open Data, tel S.Chignard, à nos enquêtes de terrain en passant par les pratiques professionnelles des réutilisateurs que nous avons étudiées, tout concorde en ce sens.

6 Conclusion : L'Open Data reconsidérée ?

Au fil de ce premier chapitre nous avons cherché à montrer comment les idéaux de l'ouverture des données se sont réalisés dans les différents projets et portails des métropoles françaises jusqu'à former au fil du temps des écosystèmes plutôt spécifiques qui témoignent des objectifs des acteurs dominants. Sans entrer ici dans le détail de chacune des métropoles étudiées dans l'ANR *OpenSensingCity*, nous avons déjà remarqué que ce processus long et complexe d'ouverture s'est joué en plusieurs étapes, au gré des injonctions législatives mais aussi de la réflexion et de l'expérience des métropoles. En effet, après plusieurs versions et dénominations les portails métropolitains Open Data se stabilisent et s'organisent à la croisée des idéaux, des visées et des caractéristiques du territoire. Au final, se dégagent différents types de portails Open Data métropolitains qui matérialisent une appropriation particulière de l'Open Data, en termes d'industrialisation de la donnée comme nous l'avons montré. Autrement dit, les types de portail stigmatisent autant l'acteur dominant, public et territorial, au sein de son réseau de partenaires, que l'utilisateur cible de l'ouverture, qu'il soit citoyen ou professionnel de la data, car il est avant tout perçu comme le réutilisateur des données rendues disponibles.

La réutilisation reste un des termes forts et caractéristiques de l'ouverture et figure d'ailleurs dans la plupart des définitions de l'ouverture, elle constitue un objectif difficile à atteindre [42] car elle relève du registre de l'idéologie et doit être reliée à deux grandes utopies contemporaines du Web (Ibekwé et Paquienséguy : 2015, 15) : l'ouverture et la participation, qui conduisent à penser que le lien se créerait de façon automatique entre offre et usages, entre données ouvertes et réutilisations, alors que ces processus doivent se construire et les pratiques professionnelles coloniser ces nouveaux espaces matérialisés par les portails Open Data jusqu'à bâtir progressivement un cadre de référence, au sens de P. Flichy (1995), porteur de l'écosystème de la réutilisation, finalement plus important et sans doute bien plus inaccessible que celui de la simple ouverture. D'ailleurs le projet public d'ouverture des données institue leur

42. Cette hypothèse a particulièrement été travaillée dans : Paquienséguy F., Dymytrova V. « 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. Disponible à l'adresse : http ://ojs.imodev.org/index.php/RIGO/article/view/204.

réutilisation comme l'objectif et la raison d'être de la mise à disposition, entre autres au travers de la Charte du G8 pour l'Ouverture des Données Publiques [43], qui voit dans la réutilisation, la possibilité d'« *acquérir des connaissances, 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* » [44]. Comme après une phase d'idéation, en termes d'ouverture, les métropoles ont dû à la fois inventer et produire, en termes de standards techniques certes, mais surtout de compétences, de gouvernance, de métiers en vue d'un écosystème informationnel. Autrement dit, tenter une structuration en termes de gouvernance et de modèle économique (Paquienséguy : 2016) témoigne d'une volonté de penser un cadre d'usages (de réutilisation) stable dans un cadre idéologique, ainsi défini par les États membres du G8 : ces 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* [45].

Le lien entre ouverture et réutilisation s'est progressivement construit depuis les premières versions des portails et les premiers services métropolitains en charge de l'ouverture et du portail ; c'est cette double structuration, revisitée au fur et à mesure de l'expérience acquise, qui a finalement conduit à l'émergence de trois types de portails. Ces types de portail que nous avons dégagés dans l'ANR s'entendent au regard de l'intermédiation qu'ils valorisent : certains portails s'adressent surtout aux data journalistes ou *data scientists* et producteurs d'applications, d'autres gardent pour cible première le citoyen et tentent une vulgarisation de l'accès et de l'utilisation des données, et enfin les derniers cherchent à valoriser les applications développées par les professionnels de la donnée à destination du grand public en les hébergeant.

Outil central de la panoplie (Labelle : 2007) de l'ouverture des données développée par les métropoles, le portail fonctionne comme le guichet unique d'accès aux données d'un territoire puisqu'il centralise les données locales de la façon la plus complète possible en cherchant à y agréger toutes les données métropolitaines, et parfois même celles qui ne sont pas soumises à l'injonction publique. Dans le même élan, il prolonge les choix politiques locaux en cherchant à assurer un lieu d'intermédiation entre producteurs de données, collectivités territoriales et différentes catégories de réutilisateurs de données. La position tenue par les portails est donc d'assurer la fonction centrale de l'entremise afin de favoriser la réutilisation. Cependant, les voies empruntées par les métropoles pour faire du courtage informationnel entre les parties productrices et réutilisatrices des données, semblent parfois aller à l'encontre des idéaux mêmes de l'ouverture.

43. 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.

44. *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.

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

Bibliographie

Arruabarrena, B. (2015, Février). Datavisualisation : des données à la connaissance. *I2D – Information, données & documents*, 52, p. 28-29.

Béraud, Ph. et Cormerais, F. (2012, Février). Industries créatives, industries de territoire : la géographie culturelle d'une métropole régionale. *Communication & Organisation*, 42, p. 203-216.

Bouquillion, Ph.(2012, Septembre). Concentration, financiarisation et relations entre les industries de la culture et industries de la communication. *Revue française des sciences de l'information et de la communication* [En ligne], mis en ligne le 01 septembre 2012, consulté le 06 juin 2019. URL : http ://journals.openedition.org/rfsic/94 .

Brookes, K. (2018, Février). Grenoble et la « Smart City ». Entretien avec Éric Piolle, maire de Grenoble

Broudoux, E. (dir.) et Chartron, Gh. (dir.). (2015). Big Data – Open Data : Quelles valeurs ? Quels enjeux ? Bruxelles : De Boeck Supérieur.

Charles, E. et Thouément, H. (2018). *La métropolisation : une opportunité pour le développement des territoires, Paris : L'Harmattan.*

Chignard, S. et Benyayer, D.(2015). Datanomics, les nouveaux business models des données, Paris : FYP Éditions

Chopplet, M. (2018). Smart City : quelle intelligence pour quelle action ? Les concepts de John Dewey, scalpels de la ville intelligente. Quaderni, 96, p. 71-86.

Dymytrova, V., Larroche, V., Paquienséguy, F. (2018). Cadres d'usage des données par des développeurs, des data scientists et des data journalistes. Livrable 3. [Rapport de recherche] EA 4147 Elico, https ://hal.archives-ouvertes.fr/hal-01730820v1.

Ibekwé, F., Paquienséguy, F. « Open, Big, Collaboration : trois utopies de l'innovation au 21e siècle », in Big Data – Open Data : Quelles valeurs ? Quels enjeux ? in Broudoux, Evelyne. Chartron Ghislaine (dir.), Louvain-la-Neuve, De Boeck Supérieur, 2015, p. 15.

Mabi, Cl.(2015, Mai). La plate-forme « data.gouv.fr » ou l'open data à la française. Informations sociales,19, p. 52-59.

Mahé, Th. et Ennassiri, K. (2014). L'open *Data a-t-il seulement un impact économique ?* https ://archinfo24.hypotheses.org/1887

Moatti, S. (2015). Les promesses de l'open data. *Alternatives Économiques*, 313, p. 38-38.

Mœglin, P. (dir.). (2016). *Industrialiser l'éducation. Anthologie commentée (1913-2012)*, Saint-Denis : Presses universitaires de Vincennes, coll. Médias

Moritz, M. (2016). Big data et villes éco-intelligentes. *I2D – Information, données & documents*, 53, p. 62-63.

Naegelen, P. (2016). Vers un domaine public des données ? *I2D – Information, données & documents*, 53, p. 25-26.

Paquienséguy, F., Dymytrova, V. (2017). Analyse de portails métropolitains de données ouvertes à l'échelle internationale. Livrable 1. 2. [Rapport de recherche]. Elico (Équipe d'accueil lyonnaise en Sciences de l'information et de la communication).hal-01449348.

Paquienséguy, F., Larroche, V., Peyrelong, M.-F. Vila-Raimondi, M., Dymytrova, V. (2016). Synthèse des résultats de l'enquête auprès de ré-utilisateurs de données ouvertes. Livrable 1. Elico. (https ://hal.archives-ouvertes.fr/hal-01432124).

Paquienséguy F., *Smart city & open data : à qui profitent les données ouvertes ?* Colloque international des Sciences du Territoires 2016, Proceedings « En quête de territoire (s), Looking for territories », Grenoble, 17-18 mars 2016, p. 351-357 https ://hal.archives-ouvertes.fr/hal-01353651.

Pollock, R., *The Open Revolution, rewriting the rules of the information age*, https ://openrevolution.net

Ricœur, P. (1984). L'idéologie et l'utopie : deux expressions de l'imaginaire social. *Autres Temps. Les cahiers du christianisme social.* 2, p. 53-64.

Teresi, L. (2011). Exploitation des données publiques : le renouveau ? *Légicom*, 47, p. 51-63.

Tétu, J.-F. (2017). P. Mœglin, dir, Industrialiser l'éducation. Anthologie commentée (1913-2012). Questions de communication, 31

Rochet, Jean-Charles et Tirole, Jean (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, 4, p. 990-1029

Verlaet, L. (2015). La deuxième révolution des systèmes d'information : vers le constructivisme numérique. *Hermès, La Revue*, 71, p. 249-254.