

OPEN DATA

Synthèse bibliographique

octobre 2013

Le mouvement d'ouverture des données publiques, plus communément appelé open data, est en plein essor depuis le milieu des années 2000. Ce mouvement, porté à l'origine par la société civile, a trouvé écho auprès des pouvoirs publics : la plupart des grandes démocraties ont modifié ou envisagent de modifier leur droit interne en vue d'une libéralisation des données accrue et mieux délimitée juridiquement. Ces dernières années, la France, sous l'impulsion de l'Europe, a ainsi largement redéfini le cadre fixé par la "loi CADA" du 17 juillet 1978 relative au droit d'accès aux documents administratifs.

Les évolutions en cours du corpus juridique relatif à l'ouverture des données, et notamment du droit européen, vont continuer à accompagner la volonté citoyenne d'accès à l'information publique mais également faciliter les réutilisations. Comment expliquer le choix d'une telle démocratisation de la donnée publique ? Quelles sont les retombées attendues et sous quelles conditions peut-on espérer une réelle plus-value de l'open data tout en limitant les risques ? La présente synthèse apporte des éléments de réponse à ces problématiques à partir d'une analyse de la bibliographie récente mais aussi des sites spécialisés et articles diffusés sur internet.

UN MOUVEMENT CITOYEN RELAYÉ PAR LES POUVOIRS PUBLICS

Définition de la donnée

La donnée peut être définie comme "*l'information brute recueillie par un capteur ou générée par l'utilisateur lui même* » selon Sylvie Servigne, maître de conférence au LIRIS de Lyon¹. Certains considèrent qu'aujourd'hui, tout le monde peut produire de la donnée et que "*les données sont partout, et de tout ordre. Tout est donnée* » [Chignard, 2012].

La culture de la collecte de la donnée n'est pas récente, elle est inhérente au bon fonctionnement de toute société. Toutefois le recueil des données et leur capitalisation ont été facilités et de ce fait amplifiés par les nouvelles technologies². D'autre part, les exigences réglementaires poussent les acteurs publics à se doter de bases de données solides leur permettant de prendre les bonnes décisions. Ainsi, Thierry Lagarde, responsable du service Observation et valorisation des données à la Délégation générale au Développement urbain du Grand Lyon souligne que les lois Grenelle "*obligent [les collectivités locales] à mettre en place des outils pour répondre aux contraintes réglementaires : avant de travailler sur un territoire, une politique d'aménagement urbain, il faut se doter d'outils de connaissance, notamment sur la biodiversité, la qualité de l'air, le bruit* »³.

Les données publiques, que l'on peut définir comme les données détenues par les organismes publics et ne relevant pas de la vie privée ou de la sécurité nationale [Lecko, 2012], sont donc de plus en plus nombreuses. Or, cette source inépuisable d'informations est collectée grâce à l'impôt : pour les militants de l'open data, il est par conséquent naturel de restituer ces données au citoyen, qui a financé leur collecte. Notons que l'on parle aujourd'hui principalement d'ouverture de la donnée publique car les démarches open data sont, en France notamment, essentiellement portées par les acteurs publics. En outre, la législation régissant ce domaine ne contraint que les organismes publics. Il faut toutefois garder à l'esprit que le mouvement open data visait à l'origine tous les détenteurs de données et que les initiatives allant dans le sens d'une libéralisation des données des acteurs privés se multiplient [Décider ensemble, 2012].

1 <http://www.millenaire3.com/Donnees-services.1387.0.html>

2 Certains chercheurs considèrent que dès 2007, 94% de notre "mémoire" était devenue numérique. Les estimations récentes de la masse totale d'informations au niveau mondial avoisinent les 295 exabytes avec un taux de croissance de plus de 50% par an [Hilbert M., López P., 2011]. En parallèle, notre capacité à stocker les données a augmenté de manière exponentielle : par exemple, un disque dur acheté 600 dollars pourrait stocker toute la musique du monde [MacKinsey Institute, 2011].

3 <http://www.millenaire3.com/Donnees-services.1387.0.html>

Les origines et le fondement du mouvement d'ouverture des données publiques

Si le principe de restitution de la donnée aux contribuables pourrait constituer l'un des arguments majeurs en faveur de la libéralisation des données publiques, il n'est pas mis systématiquement en avant par les personnes ou organismes faisant la promotion de l'open data. On parle davantage de transparence, de nouveau mode de partage favorisant l'innovation, ou encore de coproduction. Dans les années 60, la volonté citoyenne d'accéder à la donnée répondait à un besoin de s'informer, de comprendre voire de contester l'action publique [Décider ensemble, 2012]. Elle donnera plus tard naissance au mouvement open data, mieux organisé et aux objectifs plus larges, faisant écho aux évolutions sociétales, notamment à la multiplication des données et aux nouvelles perspectives offertes par l'air du numérique.

Ce mouvement prend de l'ampleur au milieu des années 2000 et incite les gouvernements à accompagner l'ouverture des données publiques. Ainsi, la plateforme gouvernementale de partage des données "data.gov" est lancée en 2009 par le Président Obama. Au Royaume-Uni, "data.gov.uk" suivra en 2010, prenant le relais de différentes initiatives de la société civile tel que "Free our data" du journal "The Guardian". En France, les collectivités locales sont les premières à créer leur plateforme (Rennes en 2010). L'Etat lance "data.gouv.fr" fin 2011.

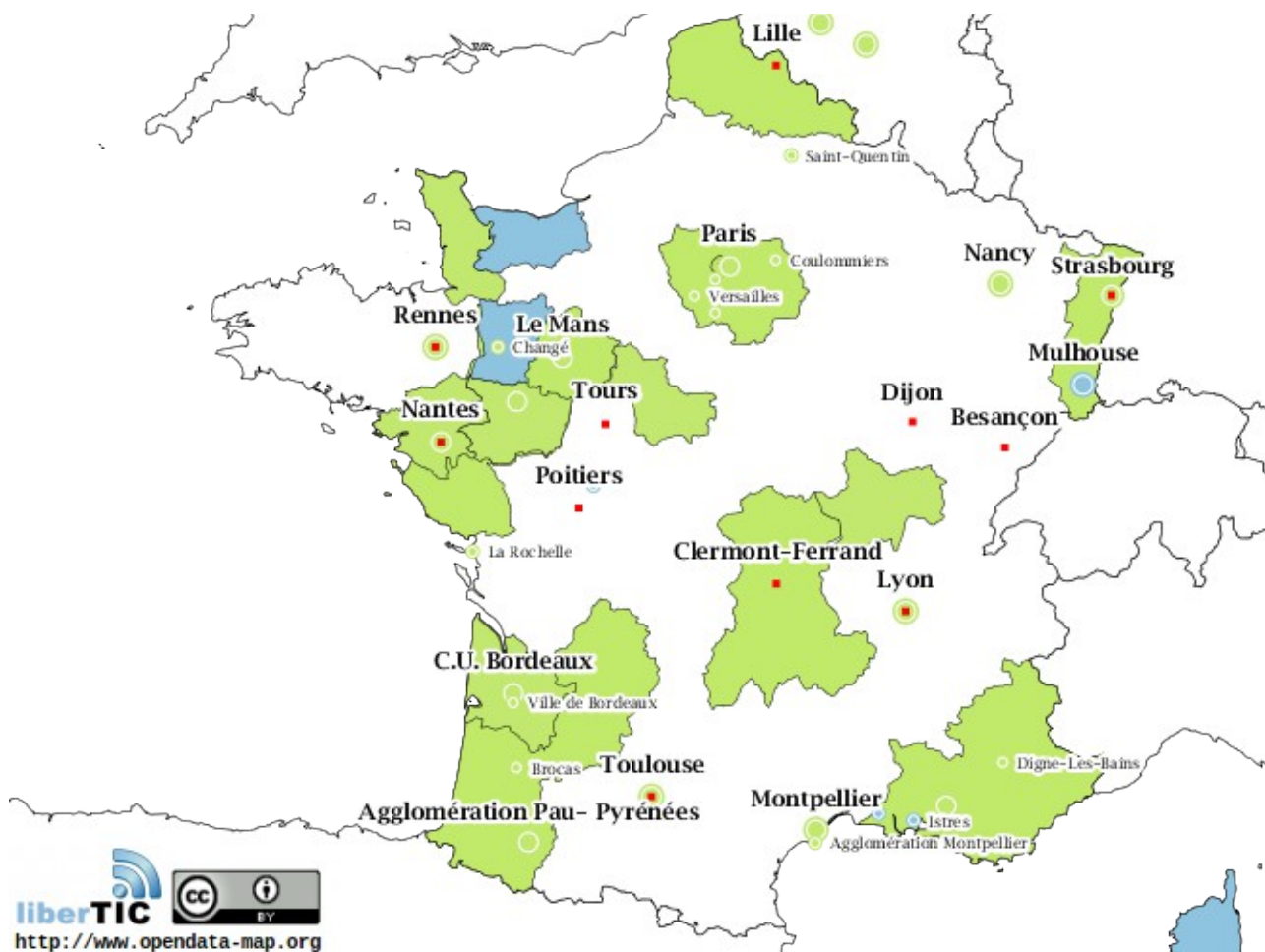


Illustration 1: Carte de France de l'open data. Source : Libertic, 2013

En vert, les territoires sur lesquels il existe des plateformes de partage des données ; en bleu, les projets en cours ; en rouge, les mouvements citoyens.

L'accompagnement du mouvement open data par les pouvoirs publics

Dans la plupart des grandes démocraties, les pouvoirs publics répondent donc à cette attente de libéralisation des données par la création de portails internet de partage mais aussi en faisant évoluer leur droit interne.

En France, le principe d'ouverture des données publiques est fondé sur l'article 1er de la loi n°78-753 du 17 juillet 1978 relative au droit d'accès aux documents administratifs. Ces derniers recouvrent tous "*les documents produits ou reçus, dans le cadre de leur mission de service public*", par les personnes de droit public. Cette loi donne la possibilité au citoyen de demander directement ces documents et leur octroie le droit de se retourner, en cas de refus, vers la Commission d'accès aux documents administratifs (CADA). Certains documents ne sont pas communicables tels que ceux relatifs au secret-défense ou ceux contenant des données à caractère personnel.

Par la suite, différents textes de loi sont venus renforcer les obligations faites aux personnes publiques de rendre leurs données accessibles. La directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public marque l'une des évolutions majeures en la matière. La directive vient poser les bases d'une réutilisation libre des informations publiques, définies comme les informations détenues dans les documents administratifs [CNN, 2012]. Ce droit à la réutilisation des données publiques est transposé en droit interne par l'ordonnance du 6 juin 2005 et le décret du 30 décembre 2005 qui modifient et complètent la loi CADA de 1978. De surcroît, l'ordonnance de 2005 rend le droit à la réutilisation opposable, la CADA étant chargée de veiller à son application.

Plus récemment, le décret n° 2011-194 du 21 février 2011 crée la mission Etalab, chargée de la mise en place d'un portail unique interministériel des données publiques (www.data.gouv.fr) ; Etalab est aujourd'hui sous la coupelle du Secrétariat général pour la modernisation de l'action publique (SGMAP) créé fin 2012 par le gouvernement. Le décret n° 2011-577 du 26 mai 2011 relatif à la réutilisation des informations publiques détenues par l'État et ses établissements publics administratifs vient quant à lui consacrer le principe de la gratuité de la mise à disposition des données, l'exception payante devant être justifiée. Jérôme Filippini, directeur du SGMAP, estime en effet, dans une interview publiée le 9 novembre 2012, que "*la valeur créée par la libération des données est infiniment plus importante que les éventuels revenus que l'Etat pourrait tirer de leur vente*"⁴.

On est donc passé progressivement d'un droit démocratique à la transparence, posé en 1978 par le législateur avec le principe d'accès aux données, à un droit économique et social de réutilisation des données, les pouvoirs publics considérant "*comme [leur] devoir de distribuer la capacité de faire des usages pertinents de [leurs] propres données*" [FING, 2011].

Cette évolution du droit français en faveur de l'ouverture et la réutilisation des données publiques va se poursuivre avec la transposition d'ici 2015 de la nouvelle directive européenne sur la réutilisation des informations du secteur public adoptée par le Parlement européen le 13 juin 2013. Cette directive réaffirme plusieurs grands principes tels que la gratuité des données et le droit à la réutilisation qui s'étendra désormais à toutes les données du secteur public ne faisant pas l'objet d'une dérogation : la directive de 2003 laissait en effet la possibilité aux personnes publiques de décider quelles informations elles souhaitaient rendre réutilisables ou non. Le périmètre d'application de ces principes couvre désormais toutes les données publiques et plus seulement celles ayant déjà été publiées. La directive vise également à favoriser la mise à disposition des données du secteur public dans des formats ouverts lisibles par machine mais elle n'impose pas les formats libres. In fine, le droit français n'est impacté qu'à la marge par cette directive puisqu'il impose déjà la plupart des obligations fixées par cette dernière, avec notamment la reconnaissance dans la loi CADA modifiée d'un droit à la réutilisation des informations publiques dès lors que les documents qui les contiennent sont accessibles [Ecoter, 2013].

Toujours concernant les évolutions à venir, l'impact de "l'acte III de la décentralisation", en lecture devant le Parlement, sur l'ouverture des données au sein des collectivités territoriales est encore difficile à évaluer. On semble s'orienter vers des obligations accrues faites aux collectivités en matière de partage de données et de possibilités de réutilisation. Toutefois, les articles du projet de loi relatifs à l'ouverture des données restent ambigus et seuls les débats parlementaires puis les décrets d'application pourront nous éclairer sur les apports

4 http://www.lesechos.fr/09/11/2012/LesEchos/21309-141-ECH_open-data---le-gouvernement-tente-de-rassurer-le-secteur.htm

réels de ce texte.

Notons enfin qu'en matière de mobilité, l'ouverture des données en cours permet de remplir les objectifs fixés par la directive européenne 2007/2/CE du 14 mars 2007, dite directive Inspire [AFIMB, 2013]. Cette directive vise à établir une infrastructure d'information géographique dans la Communauté européenne pour favoriser la protection de l'environnement. Il s'agit donc de mettre en place "*un ensemble de services d'information disponibles sur Internet, répartis sur les sites web des différents acteurs concernés, et permettant la diffusion et le partage de données géographiques*"⁵. Cette directive s'applique aux données géographiques détenues par les autorités publiques, dès lors que ces données sont sous forme électronique et qu'elles concernent l'un des 34 thèmes identifiés, parmi lesquels figurent les réseaux de transport.

Une évolution des règles de droit à poursuivre

L'évolution décrite ci-avant montre que l'open data gagne peu à peu du terrain, sous l'impulsion de la société civile et avec l'appui du législateur qui a rendu plus aisé l'accès et la réutilisation des données. Toutefois, il reste des barrières législatives à lever.

On peut notamment se demander si un assouplissement des règles relatives à la réutilisation des données culturelles ne serait pas profitable au citoyen mais aussi au patrimoine, offrant à ce dernier une visibilité accrue. Cet assouplissement n'est réalisable que sous la condition d'une clarification des règles de la propriété intellectuelle.

En outre, pour une partie de la doctrine, la loi CADA modifiée est encore trop restrictive quant à l'obligation de mise en ligne des données publiques. Si le citoyen peut demander à avoir accès à la plupart des données publiques (des exceptions liées par exemple au secret défense ou à la vie privée étant prévues par la loi), les administrations ne sont pas contraintes à la publication de l'ensemble de ces données : quand la donnée n'est pas mise en ligne, c'est à l'administré de formuler une demande d'accès, la procédure pouvant prendre plusieurs mois. La logique qui prévaut est donc encore celle de la demande et non celle de l'offre. Or, dans une optique de partage et de promotion de l'innovation, les acteurs publics devraient tendre vers une logique d'offre afin de faciliter les réutilisations et donc l'émergence de nouveaux services [CNN, 2012].

Enfin, certains militent pour un élargissement du périmètre des informations publiques. D'une part, on peut regretter que la définition actuelle des documents administratifs et des informations publiques exclut de nombreuses données. Si beaucoup d'exceptions se justifient (données relatives au secret défense, au secret bancaire ou statistique, données à caractère personnel...), d'autres pourraient être aménagées. D'autre part, des données pour lesquelles le potentiel d'innovation de l'open data est très fort sont placées par le législateur en-dehors du champ du droit à la réutilisation. C'est notamment le cas des données collectées et produites dans le cadre d'une mission de service public à caractère industriel ou commercial. De nombreuses données relatives au transport sont ainsi accessibles mais non réutilisables [CNN, 2012]. Pourtant, les développeurs et le grand public manifestent un intérêt certain pour ces données, la mobilité étant l'un des domaines où les perspectives d'innovation sont les plus fortes.

⁵ <http://www.developpement-durable.gouv.fr/La-directive-europeenne-Inspire-de.html>

LES NOUVELLES PERSPECTIVES OFFERTES PAR L'OPEN DATA

Les bénéfices de l'open data pour la collectivité peuvent être nombreux et profitables à tous, acteurs publics, secteur privé ou citoyens. Nous dirons d'abord quelques mots des apports attendus de la libéralisation des données en matière de mobilité, thème de l'étude commandée par le CERTU dans laquelle s'inscrit la présente synthèse bibliographique, avant d'aborder plus globalement le nouveau champ des possibles ouvert par l'open data.

Les apports de l'open data spécifiques au domaine de la mobilité

Les transports et plus globalement la mobilité figurent parmi les thèmes phares des plateformes d'ouverture des données mises en place par les acteurs publics. Les données sur les transports sont ainsi présentes sur 51% des plateformes enquêtées dans le cadre d'une étude récente (illustration ci-après).

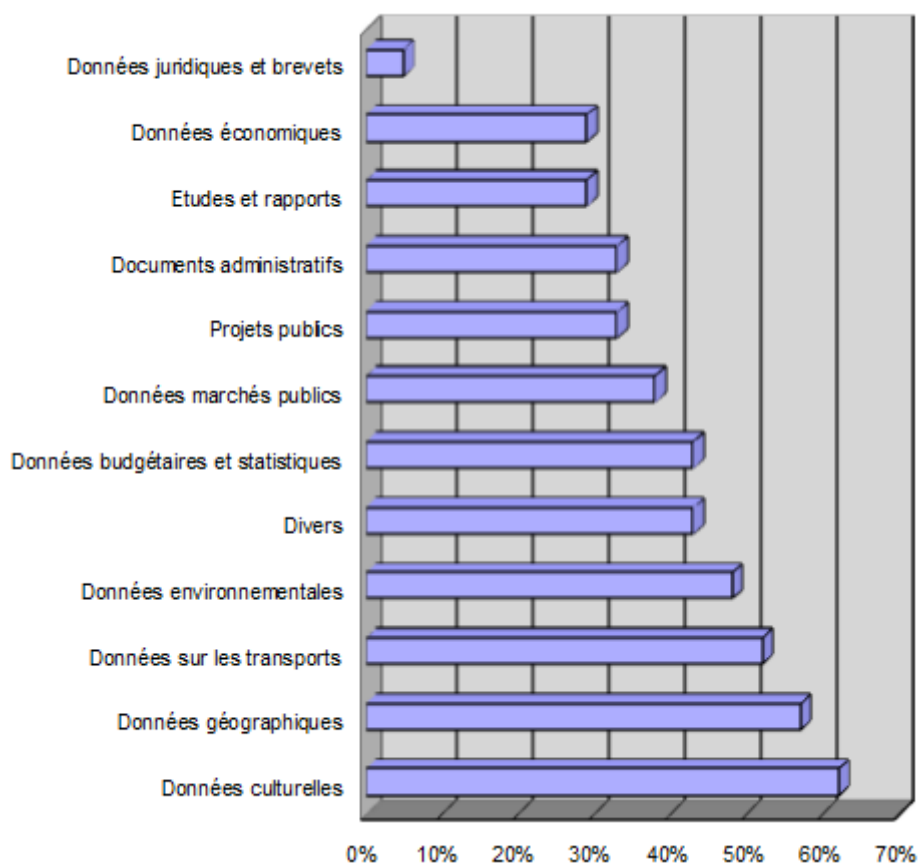


Illustration 2: types de données publiées en ligne par 57 organismes producteurs de données publiques enquêtés. Source [Serdalab, 2012]

Ces derniers répondent à l'attente forte des citoyens et des développeurs pour la mise en ligne d'informations relatives aux déplacements : les premiers veulent des données qui leur permettent de se rendre plus facilement d'un point A à un point B, dans un monde où il faut être de plus en plus mobile pour ne pas être exclu socialement ; les seconds veulent accéder à des bases qui leur donneront la possibilité de créer des services innovants répondant aux besoins croissants de la population en matière de mobilité. Ces services ne se limitent pas aux applications mobiles, certains développeurs proposant par exemple de l'information voyageur via une montre ou une "box" posée sur la table du salon⁶.

Par la libéralisation des données publiques relatives à la mobilité, l'open data peut favoriser l'émergence de services que les acteurs publics n'auraient pu mettre en place seuls, faute de temps ou de compétences, mais

6 <http://donneesouvertes.info/>

aussi car la mobilité est un domaine où la réactivité, la mise à jour des données, est essentielle si l'on veut proposer des services pertinents. Or, cet objectif de réactivité peut être atteint grâce à la remontée des informations par les voyageurs eux-mêmes, permettant une actualisation en temps réel des bases de données : en matière de mobilité, plus encore peut-être que dans les autres domaines, il faut donc partager avec les citoyens et savoir utiliser la ressource voyageur pour améliorer les bases et les services qui en découlent.

Les apports transversaux de l'open data

Si les enjeux spécifiques à la mobilité sont réels, les apports de l'open data pour la collectivité peuvent être globaux et transversaux.

L'un des objectifs les plus cités dans les écrits analysés est celui de transparence. L'open data doit permettre de rendre l'action publique moins opaque et donner un droit de regard au citoyen. La feuille de route du gouvernement en matière d'ouverture et de partage des données publiques⁷, publiée le 28 février 2013, propose ainsi de "*renforcer la confiance des citoyens grâce à une plus grande transparence de l'action publique (par exemple, avec la mise en ligne d'informations sur les comptes publics, la pollution ou les délits constatés)*". La possibilité pour tout citoyen d'accéder aux données renforce la légitimité du pouvoir en place et donc la démocratie. De plus, soumettre un travail à la critique pousse à en améliorer la qualité : "*si les citoyens disposent, par exemple, de la performance des hôpitaux acte par acte, les établissements de santé pourraient être incités à se renforcer dans les actes où ils sont les moins bons*"⁸.

Des effets sur l'attractivité du territoire "couvert" par l'ouverture des données peuvent également être escomptés. L'open data permettra par exemple aux professionnels du tourisme de proposer des services plus complets et attractifs et donc d'attirer des voyageurs, aux entreprises ou aux familles désirant s'implanter ou se loger de disposer d'un panel d'informations plus large. Mais c'est la démarche de libéralisation des données en elle-même, avec son effet mobilisateur et la dynamique positive qu'elle engendre, qui constituera le premier facteur d'attractivité [FING, 2011].

Un autre apport attendu de l'open data par les collectivités locales est le décloisonnement entre leurs services en interne, et, en externe, entre les acteurs publics du territoire. La feuille de route du gouvernement souligne ainsi l'opportunité "*d'améliorer le fonctionnement administratif (comme en témoigne la très forte utilisation des données publiques par l'administration elle-même)*".

En interne, la mise en place d'une plateforme open data suppose, d'une part, de réaliser un travail de capitalisation et de contrôle de la qualité des données ce qui favorise la constitution de bases de données plus complètes, mieux construites et actualisées. D'autre part, les services concernés doivent échanger, se réunir : le personnel prend ainsi connaissance des données à disposition dans les différents services mais aussi des modes de travail des uns et des autres ou encore des difficultés rencontrées dans tel ou tel domaine. In fine, ces échanges peuvent aboutir à une amélioration de l'organisation interne de la structure publique⁹.

Plus encore, la mise en place d'une plateforme "mutualisée", référençant ou hébergeant les données de plusieurs organismes publics, va étendre le périmètre des échanges et peut-être faciliter des partenariats futurs entre personnes publiques.

Surtout, l'open data semble être perçu comme un levier majeur au service d'une collaboration plus étroite entre les pouvoirs publics, le secteur privé et les citoyens. L'ouverture des données publiques pourrait ainsi conduire à l'émergence d'un système de coproduction entre ces différents acteurs, l'administration fournissant la donnée pour que la société civile puisse innover et proposer des services utiles à la collectivité.

7 http://www.gouvernement.fr/sites/default/files/fichiers_joints/donnees-publiques.pdf

8 Propos de Gilles Babinet, ancien président du Conseil national du numérique [Décider ensemble, 2012]

9 <http://www.millenaire3.com/Charles-Nepote-Services-urbains-Open-data-et-m.122+M5f6b88832aa.0.html>

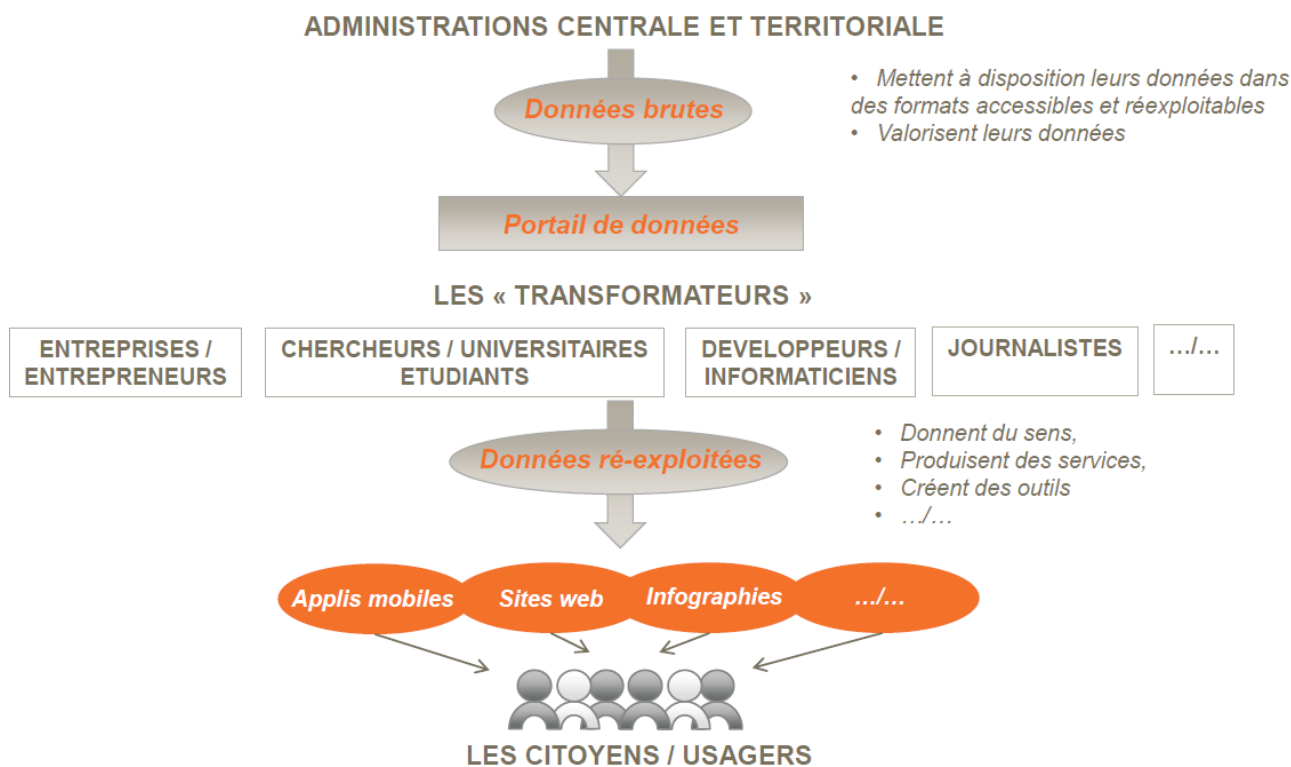


Illustration 3: source : [Lecko, 2012]

Certains parlent même de "*processus d'innovation ouverte*", qui poserait les bases d'un système itératif et qui conduirait les organismes publics à revoir leur modèle de fonctionnement pour mieux tenir compte des réactions et suggestions des citoyens." *Un acteur public a du mal à innover tout seul ; en ouvrant des données, il voit sortir des usages auxquels il ne s'attendait pas, qui participent à l'innovation du territoire*"¹⁰.

Cette coproduction peut par exemple être profitable aux petites structures publiques qui ne disposent pas des moyens humains, matériels ou des compétences pour mener à bien tel ou tel projet innovant. On peut donner l'exemple de Plouarzel, petite commune de 3150 habitants, confrontée au problème de mise à jour et d'édition de son plan communal et ne disposant pas des moyens nécessaires pour faire fonctionner un service technique : cette commune a fait le choix de mettre les cartes qu'elle avait à disposition sur OpenStreetMap, et de proposer à la population de les enrichir. Les cartes OpenStreetMap ainsi complétées par le crowdsourcing sont devenues l'outil de référence de l'Office municipal du tourisme [FING, 2011].

L'externalisation d'une partie de la "production" de la structure publique peut également laisser espérer des gains en terme de temps voire des gains économiques. Toutefois, ces derniers sont difficiles à chiffrer. En interne, les collectivités locales qui mettent en place une plateforme de partage de données peuvent réaliser des économies du fait du transfert de la production de services innovants au tissu économique local mais ces gains ne semblent pas significatifs et ne constituent pas la motivation première des acteurs publics. En revanche, les estimations laissent entrevoir des retombées économiques substantielles pour le territoire : selon une étude récente [Vickery, 2011] commandée par la Commission européenne, un accès facilité aux données pourrait augmenter de 40 milliards d'euros le marché européen des informations du secteur public, aujourd'hui estimé à 28 milliards d'euros par an. L'impact économique global, direct et indirect, atteindrait 140 milliards d'euros par an. La ville de Washington, quant à elle, a évalué à 4000 % le retour sur investissement suite à l'organisation d'un concours de développement d'applications sur la base de données publiques ouvertes [Lecko, 2012]. En mettant à disposition des données qui permettront de générer des services, l'open data peut contribuer à créer de l'activité sur le territoire et ainsi stimuler le tissu économique.

¹⁰ <http://www.millenaire3.com/Charles-Nepote-Services-urbains-Open-data-et-m.122+M5f6b88832aa.0.html>

Par les liens qu'il crée et les opportunités qu'il offre, l'open data est donc source de plus-values importantes pour les acteurs publics, pour les entreprises et pour le citoyen. Toutefois, il reste actuellement de nombreux freins à lever pour réaliser pleinement ce potentiel.

LES OBSTACLES ACTUELS À L'EXPLORATION DU CHAMP DES POSSIBLES OUVERT PAR L'OPEN DATA

L'analyse des publications relatives à l'open data met rapidement en lumière des obstacles majeurs au développement du mouvement de libéralisation des données ainsi qu'à l'optimisation des possibilités de réutilisation de ces dernières.

Les obstacles internes à la structure souhaitant libéraliser ses données

La première barrière à franchir est souvent interne à la structure souhaitant mettre ses données en partage. Les agents au sein de l'organisme public concerné peuvent opposer une résistance au changement important que représente l'ouverture des données au grand public. Cette résistance est souvent liée à la crainte d'être jugé sur la qualité voire l'exactitude de la donnée produite. Il appartient alors à la hiérarchie d'informer sur la démarche et les objectifs poursuivis, l'amélioration de la qualité des données grâce aux "remontées" des utilisateurs de la plateforme de partage faisant généralement partie de ces objectifs sans pour autant viser à "fliquer" les agents¹¹.

En outre, les décideurs au sein de l'organisme peuvent eux-même hésiter à se lancer dans l'open data. D'une part, on peut regretter que *"la détention de l'information [reste] un privilège sur lequel beaucoup fondent leur pouvoir; qu'il n'est pas naturel pour eux de partager"*, comme le notait Séverin Naudet au moment de quitter ses fonctions à la tête d'Etalab. Cet Etat conservateur, adepte du secret, très centralisé, était également pointé du doigt par Gilles Babinet, ex-patron du Conseil national du numérique¹². Mais cette culture du secret ne peut caractériser l'ensemble des décideurs publics. Pour la plupart, les motifs d'hésitation résideront dans le travail de capitalisation, de vérification et d'actualisation des données à réaliser, donc dans le temps à consacrer et les coûts associés. Le coût "technique" est lié notamment au développement des API¹³, à l'acquisition éventuelle d'un serveur puissant pour répondre au trafic induit par les requêtes aux API, au renforcement des protections contre d'éventuelles attaques du réseau. Il faut également tenir compte des coûts liés au temps passé par les agents de la structure publique pour faire fonctionner la plateforme, compléter ou actualiser les bases de données voire échanger avec les utilisateurs. Certains jugent que les coûts sont faibles par rapport à ceux généralement pratiqués par les consultants pour le développement de logiciels ou par rapport à ceux engendrés par une démarche purement interne sans partage de la production avec la société civile [Kaufman, Wagner, 2012]. D'autres rappellent que la "facture" n'est pas neutre. Le budget moyen s'élèverait à 83 000 euros de ressources humaines, hors informatique [Serdalab, 2012]. Mais les estimations du coût global de mise en place et de suivi d'une plateforme open data sont très variables, allant de quelques dizaines de milliers à plusieurs centaines de milliers d'euros¹⁴. Pour faire face à ces nouvelles dépenses à consentir dans un contexte de diminution des ressources publiques, les acteurs publics

11 <http://www.millenaire3.com/Charles-Nepote-Services-urbains-Open-data-et-m.122+M5f6b88832aa.0.html>

12 <http://www.acteurspublics.com/2013/02/12/1-open-data-pourquoi-ca-coince>

13 Définition Wikipédia : "une interface de programmation (abr. *API* pour *Application Programming Interface*) est un ensemble normalisé de classes, de méthodes ou de fonctions qui sert de façade par laquelle un logiciel offre des services à d'autres logiciels. Elle est offerte par une bibliothèque logicielle ou un service web, le plus souvent accompagnée d'une description qui spécifie comment des programmes *consommateurs* peuvent se servir des fonctionnalités du programme *fournisseur*".

14 <http://www.lagazettedescommunes.com/184790/open-data-quels-couts-pour-les-collectivites-territoriales/>

pourraient être tentés de monétiser les données. Si le décret n° 2011-577 du 26 mai 2011 précité est venu fortement contraindre les possibilités d'une telle monétisation, certains décideurs publics ne seraient pas contre une évolution du droit sur ce point : *"si l'on veut inciter les administrations à participer au mouvement de libération des données, il faut certaines contreparties* », rappelait ainsi fin 2012 Olivier Schrameck, conseiller d'Etat et membre de la Commission sur la rénovation et la déontologie de la vie publique¹⁵.

ZOOM sur la problématique spécifique de l'ouverture des données des entreprises

On parle essentiellement aujourd'hui d'ouverture des données publiques, mais, comme cela a été rappelé précédemment, le mouvement open data a un objectif plus large, visant à la libéralisation de l'ensemble des données, publiques et privées.

En France, si aucun texte de loi n'oblige aujourd'hui les entreprises à mettre en partage leurs données privées, les opportunités offertes par l'open data semblent séduire peu à peu les responsables des grands groupes. Une enquête réalisée en 2011 [Bluenove, 2011] révèle que seulement 4% d'entre eux déclarent ne pas être prêts à partager des données internes avec des tiers. Pour les autres, l'open data peut s'inscrire dans une stratégie d'innovation. Par exemple, 62 % des enquêtés pensent que l'ouverture des données leur permettrait de développer un écosystème de partenaires innovants. En outre, 63% estiment pouvoir améliorer la satisfaction-client grâce à l'ouverture des données. En revanche, 55% ne pensent pas pouvoir générer de nouveaux revenus grâce à l'open data.

Le principal frein à l'ouverture de ces données privées, outre les enjeux stratégiques et commerciaux pour certaines d'entre elles, semble être le manque d'information sur l'open data : seulement 12% des enquêtés disent "très bien" connaître le concept d'open data. Cette méconnaissance ne permet pas de bien cerner les enjeux de la démarche et de les partager au sein de l'entreprise.

Une mutualisation des plateformes difficile à mettre en place

Les perspectives de bénéfices liés à une monétisation des données étant aujourd'hui limitées, les structures publiques souhaitant réduire les coûts de leur démarche open data peuvent se tourner vers une autre piste : la mutualisation des plateformes avec d'autres partenaires, qui, outre la simplification de l'accès pour l'utilisateur, pourrait permettre de réaliser des économies d'échelle.

Mais cette option pose des questions d'ordre politique et stratégique pour la structure publique : il faut accepter, d'une part, le principe d'une gestion commune, donc d'une perte d'indépendance, et d'autre part, le risque de perdre en visibilité car la structure ne disposera pas de sa plateforme propre qui lui permettrait de communiquer auprès de ses administrés. Or, certaines structures publiques semblent vouloir utiliser l'open data comme un outil de communication, pour montrer aux citoyens leur volonté d'aller vers plus de transparence, plutôt que comme un vrai outil de partage : on voit donc parfois émerger des plateformes d'ouverture des données qui ne mettent en ligne que peu de jeux de données ou des jeux difficilement réutilisables, le but premier de la démarche étant de marquer la présence de l'acteur public au sein du mouvement open data. On parle alors de datawashing [Décider ensemble, 2012]. Il faut donc convaincre les décideurs publics que les bénéfices d'une vraie démarche de libéralisation des données et, plus encore, d'une mutualisation des plateformes seraient supérieurs à ceux générés par une telle stratégie de communication.

Comme dit précédemment, l'administration pourrait réaliser des économies d'échelle du fait de cette mutualisation. Mais c'est surtout le tissu économique local qui profiterait d'un tel regroupement des plateformes [AFIMB, 2013]. En effet, l'éclatement actuel des plateformes de partage de données n'offre pas aux développeurs un modèle économique pertinent. Les services créés à partir de la base de données mise en partage ne peuvent en effet être utilisés que sur le territoire couvert par ces données c'est-à-dire le territoire de la collectivité locale ayant mis en place sa plateforme. Or, pour que les développeurs puissent vivre de la commercialisation de leurs services, il faudrait que ces derniers soient utilisables à l'échelle nationale, ou pour le moins supra-régionale. La mutualisation des plateformes des collectivités locales pourrait permettre

15 http://www.lesechos.fr/16/10/2012/lesechos.fr/0202329311759_open-data---l-etat-cherche-a-monetiser-les-donnees-publiques.htm

de faciliter la création de ces services à grande échelle.

La manque d'harmonisation et de standardisation

Une autre solution serait d'harmoniser les formats, les référentiels ou encore les normes de construction des catalogues de jeux de données : actuellement, la multiplicité des formats, par exemple, rend très complexe la création d'un service basé sur la réutilisation de jeux de données hébergés par plusieurs plateformes. Ce type de réutilisation "croisée" deviendrait plus aisée avec la généralisation de standards communs aux différentes plateformes. Plusieurs mouvements en faveur de l'open data, tels qu'Opendata France, proposent ainsi des solutions pour une meilleure harmonisation¹⁶.

Ce principe d'harmonisation devra également être appliqué aux licences. En effet, lors du lancement d'une plateforme de partage de données, l'une des premières questions qui se pose concerne le choix de la licence, qui encadre le droit de réutilisation accordé aux tiers. En France, les administrations publiques utilisent principalement deux types de licence : la licence ouverte proposée par Etalab et la licence ODbL (Open Database License) de l'Open Knowledge Foundation. La première permet de diffuser et d'adapter les données gratuitement à condition de mentionner la source, le réutilisateur étant libre de choisir si ses données sont elles-mêmes réutilisables. La seconde permet d'exploiter publiquement la donnée, de manière commerciale ou non, et applique la clause de partage à l'identique : le réutilisateur doit donc choisir la même licence ce qui garantit une possibilité de réutilisation continue. Le choix de l'une ou l'autre licence apparaît donc comme stratégique : on privilégie une logique plus libérale avec la licence ouverte et plus commerciale mais aussi plus collectiviste avec ODbL.

Pour les développeurs, le fait que toutes les plateformes ne répondent pas à la même licence, donc aux mêmes règles de réutilisation, rend très complexe l'exploitation croisée de jeux de données hébergés par des plateformes différentes¹⁷.

La mauvaise qualité et le défaut d'actualisation des données

Les développeurs se heurtent à d'autres difficultés que celles liées à l'harmonisation lorsqu'ils veulent réutiliser des données mises en partage. La qualité de la donnée étant l'un des principaux obstacles rencontrés : les jeux de données ne sont pas toujours complets par exemple, ou non actualisés. A titre d'illustration, si le portail www.data.gouv.fr a eu le mérite de "*sensibiliser les administrations au processus de publication des données*" et "*de placer la France sur la carte mondiale de l'open data*", beaucoup regrettent qu'il ait privilégié "*la quantité au détriment de la qualité*" avec "*des budgets publiés en PDF, des agrégations rendant l'interprétation impossible*"¹⁸ etc...

Les données mises en ligne ne sont pas toujours les plus "stratégiques"...

En outre, les acteurs publics ne mettent pas toutes leurs informations en ligne et certaines données à fort potentiel de réutilisation échappent ainsi à la libéralisation. Il faut veiller, si l'on veut voir apparaître des services issus de la réutilisation des données publiques vraiment utiles au citoyen, à mettre en ligne des données stratégiques. En matière de mobilité, les développeurs manifestent ainsi un intérêt fort pour les données en temps réel, encore trop rarement partagées par les opérateurs de transports en commun : certains d'entre eux considèrent que ces informations sont trop sensibles [Bluenove, 2011], d'autres sont tout simplement en phase de construction de leur offre et devrait être en mesure de diffuser ces données à moyen terme.

Concernant ces données relatives à la mobilité, l'un des problèmes fréquemment soulevés relève de la propriété de la donnée : les données des réseaux de transports en commun, par exemple, appartiennent très souvent aux sociétés exploitant ces réseaux qui sont rarement prêtes à les partager de peur de perdre un

16 <http://opendatafrance.net/les-ateliers/atelier-3-dataset-des-datasets/>

17 <http://inno3.fr/blog/reflexions-juridiques-sur-lopen-data-en-france>

18 <http://www.data-publica.com/content/2012/11/loopen-data-est-tres-mal-estime-interview-de-claire-gallon-de-lassociation-libertic/>

avantage commercial par rapport à la concurrence. Les autorités organisatrices de transport (AOT) doivent donc, lorsqu'elles passent une délégation de service public avec ces exploitants, prévoir une clause ouvrant la possibilité de libéraliser ces données¹⁹.

Le déficit d'accompagnement des pouvoirs publics

Enfin, les pouvoirs publics doivent probablement mieux accompagner les utilisateurs, professionnels ou non de l'exploitation de données, afin de répondre à l'ambition initiale d'ouverture pour tous du mouvement open data. En effet, pour que la transparence soit réelle, le grand public doit pouvoir accéder facilement à la donnée, la comprendre facilement. Or, trop de plateformes sont aujourd'hui orientées vers les spécialistes de la donnée. Il faut donc prévoir plus de visualisations "simples" pour le grand public, limiter le nombre de formats proposés sur certains portails afin de ne pas trop complexifier l'accès à l'information et aider le citoyen à s'approprier les données. Cet accompagnement, qui n'entre pas forcément dans le champ des missions des pouvoirs publics, est déjà mis en place sur certains territoires comme à Nantes avec les associations LiberTIC et Médiagraph qui organisent chaque année plusieurs manifestations autour du numérique avec des ateliers ouverts à tous.

➔ **Un exemple de data-visualisation : « L'investissement étranger en Afrique », par Ivan Colic**

Foreign Investment in Africa

WHERE TO INVEST
Should there be a need to invest in Africa, the infographic provides a clear overview of the most attractive countries for investment. It also provides a clear overview of the most attractive countries for investment.

INVESTMENTS
Investment in \$ Billions

Country	Investment (\$ Billions)
Egypt	6712
Nigeria	5797
South Africa	5628
Morocco	2000
Algeria	1500
Tunisia	1000
Kenya	500
Other countries	Various smaller amounts

SOURCES:
data.worldbank.org
www.gomp.com/publicdata

MORE COOLNESS?
#INFOVIZ
afrographique.tumblr.com
Infographic designed by @ivancolic

Source : <http://expoviz.fr/>
Site web de l'auteur : <http://afrographique.tumblr.com/>

Création présentée dans le cadre de l'événement « Expoviz » de La Fonderie. Sur cette data-visualisation, les pays africains sont représentés selon une échelle correspondant aux montants des investissements étrangers qu'ils reçoivent.

Illustration 4: Extrait du rapport [Décider ensemble, 2012]

19 <http://www.millenaire3.com/Charles-Nepote-Services-urbains-Open-data-et-m.122+M5f6b88832aa.0.html>

CONCLUSION

Si le marché de la data existe bien, certains estiment que celui de l'open data est encore à définir²⁰. Il reste à trouver le bon modèle économique, celui qui satisfera les développeurs mais aussi les structures, publiques ou privées, mettant leurs données en partage. Pour y parvenir, il faut lever des freins juridiques, techniques, mais aussi politiques notamment pour parvenir à mutualiser les différentes initiatives. Toutefois, avant d'aller plus loin dans la libéralisation des données, ne doit-on pas se demander si la crainte affichée par certains élus vis-à-vis de cette gouvernance ouverte qui se dessine n'est pas fondée ? Ne va-t-on pas vers une dépossession de l'action publique, vers une concurrence entre services créés via l'open data et service public ? Il est important de garder à l'esprit ces questionnements dans l'optique d'une extension du mouvement de libéralisation des données profitable à tous.

Quoi qu'il en soit, la feuille de route du gouvernement en matière d'ouverture et de partage des données publiques²¹, dévoilée début 2013, prône une ouverture croissante. Elle préconise également un meilleur accompagnement des pouvoirs publics et souligne la nécessité de cibler les "*données stratégiques et de qualité*".

BIBLIOGRAPHIE

La littérature analysée pour rédiger cette synthèse est relativement fournie et très récente, ce qui s'explique par le fait que le véritable essor du mouvement d'ouverture des données date du milieu voire de la fin des années 2000. In fine, seuls les aspects organisationnels liés à l'open data (par exemple les modalités de mise en oeuvre d'une plateforme pour une collectivité ou la réorganisation à mettre en place en interne) ainsi que certains aspects techniques (touchant notamment aux référentiels de données cartographiques) sont peu traités. En revanche, les obstacles actuels à la mise en place d'une plateforme et les pistes pour les dépasser sont largement commentés. On peut toutefois regretter le faible nombre de rapports de recherche, en France et en Europe (nous avons axé notre analyse bibliographique principalement sur ce périmètre), sur cette thématique de l'ouverture des données. Des recherches tournant autour de la gouvernance, des interactions territoriales ou des rapports de pouvoir liés à la disponibilité des données seraient très certainement riches d'enseignements.

Liens internet utilisés:

<http://www.acteurspublics.com/>

<http://www.data-publica.com/>

<http://www.developpement-durable.gouv.fr/>

<http://donneesouvertes.info/>

<http://inno3.fr/>

<http://www.lagazettedescommunes.com/>

<http://www.lesechos.fr/>

<http://www.millenaire3.com/>

<http://opendatafrance.net/>

<http://www.opendata-map.org/>

<http://www.redsen-consulting.com/>

20 <http://www.millenaire3.com/Charles-Nepote-Services-urbains-Open-data-et-m.122+M5f6b88832aa.0.html>

21 http://www.gouvernement.fr/sites/default/files/fichiers_joints/donnees-publiques.pdf

Ouvrages analysés :

AFIMB, *L'open data dans le domaine du transport : analyse des premières initiatives et recommandations*, 2013

Cette étude porte sur le domaine des données relatives au transport public. Elle apporte une analyse des démarches open data existantes et propose aux promoteurs de ces initiatives des pistes d'amélioration pour dynamiser et surtout pérenniser les réutilisations dans des services au public. Cinq sites ont été analysés pour formuler ces recommandations : Rennes, Bordeaux et Gironde, Toulouse, Nantes, Loire Atlantique.

Bluenove, *Open data : quels enjeux et opportunités pour l'entreprise ?*, 2011

Cette étude est l'une des seules (en France pour le moins) qui traitent de manière approfondie de la problématique de l'ouverture des données pour l'entreprise.

Chignard Simon, *Open data, comprendre l'ouverture des données publiques*, Fyp éditions, 2012.

Voir le site <http://donneesouvertes.info>.

Cet ouvrage répond essentiellement à deux questions : comment concrètement initier et animer une démarche d'ouverture de données publiques ? Comment utiliser ces données pour construire de nouveaux services, pour réinventer la relation entre citoyens et administrations, pour développer l'économie d'un territoire ?

Conseil national du numérique (CNN), avis n°12 relatif à l'ouverture des données publiques, 2012

Cet avis du CNN comprend plusieurs recommandations très concrètes pour faciliter le développement de l'open data en France (réforme de la loi CADA, promotion de la réutilisation des données culturelles, accentuation de l'effort d'ouverture sur les jeux de données stratégiques à fort potentiel politique, social ou économique, mise en place d'une gouvernance des données publiques en faisant notamment évoluer les missions de la CADA etc...)

Décider ensemble, *Ouverture des données publiques et participation : quels enjeux démocratiques*, 2012

Ouvrage très complet qui revient sur l'historique de l'open data, met en lumière les problématiques actuelles et donnent quelques pistes d'évolution.

Ecoter, lettre hebdomadaire des TIC n°351, avril 2013

Cette lettre évoque notamment la révision de la directive européenne de 2003. Elle souligne que si l'instauration par cette révision d'un véritable droit de réutilisation constitue une vraie avancée au niveau européen, elle n'impacte que faiblement la France dont le droit interne fixait déjà des règles de réutilisation très profitables au citoyen.

FIG, *Guide pratique de l'ouverture des données publiques territoriales*, 2011

Le rapport de la FIG consacre tout un chapitre sur la marche à suivre par une collectivité territoriale pour réussir sa démarche d'ouverture des données : quelles étapes respecter pour engager une politique de réutilisation des données publiques, comment définir sa politique d'ouverture des données, comment rendre les données accessibles etc... ?

Hilbert M., López P., *The World's Technological Capacity to Store, Communicate, and Compute Information*, Science, 332, 60-65, 2011

Ce rapport (référéncé dans l'étude de Bluenove notamment) donne des estimations sur le phénomène du big data, qui permettent de mettre en lumière l'importance de la donnée dans le monde d'aujourd'hui.

Kaufman Sarah M., Wagner Robert F., *Getting started with open data : a guide for transportation agencies*, 2012

Ce rapport évoque les différents postes de dépenses liés au montage et au développement d'une démarche open data. Il indique que certains décideurs estiment que les coûts sont faibles par rapport à ceux généralement pratiqués par les consultants pour le développement de logiciels ou par rapport à ceux engendrés par une démarche conduite entièrement en interne sans co-production avec la société civile.

Lecko, *L'Open Data au coeur de la nouvelle relation entre collectivités, entreprises et citoyens*, 2012

Cette étude rappelle dans une première partie les objectifs de l'open data et les grandes problématiques liées à l'ouverture des données. Elle présente dans une deuxième partie une analyse des principaux projets open data français et

internationaux. Enfin, elle fait le point sur les pistes de réflexion que doivent engager les collectivités souhaitant se lancer dans un projet open data.

McKinsey Institute, *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, 2011

Ce rapport (référéncé dans l'étude de Bluenove notamment) souligne notamment les progrès technologiques réalisés ces dernières décennies qui permettent le stockage de données de plus en plus nombreuses et qui facilitent donc la constitution et le partage de bases de données.

Markess International, *Clés de succès d'une gestion intelligente des territoires avec le numérique : référentiel des pratiques*, 2012

A partir d'une enquête auprès de plus de cent "décideurs" au sein des collectivités locales françaises, cette étude présente des statistiques sur le potentiel d'évolution de l'open data en France dans les prochaines années.

Serdalab, *Open data en France : acteurs, projets et tendances*, 2012

Cette étude donne notamment des estimations de coûts pour le montage et l'exploitation d'une plateforme open data.

Vickery Graham, *Review of recent studies on PSI re-use and related market developments*, 2011

Le "rapport Vickery", commandé par la Commission européenne, donne des estimation des retombées économiques potentielles de l'open data. Les chiffres avancés sont très contestés par la communauté de l'open data.

Synthèse réalisée par : Nicolas JOUVE, CETE Nord-Picardie (Direction territoriale Nord-Picardie du CEREMA)