

L'Open Data en collectivité à la lumière des données de mobilité



L'Open Data en collectivité à la lumière des données de mobilité

Mars 2015

Collection Connaissances

Cette collection présente l'état des connaissances à un moment donné et délivre de l'information sur un sujet, sans pour autant prétendre à l'exhaustivité. Elle offre une mise à jour des savoirs et pratiques professionnelles incluant de nouvelles approches techniques ou méthodologiques. Elle s'adresse à des professionnels souhaitant maintenir et approfondir leurs connaissances sur des domaines techniques en évolution constante. Les éléments présentés peuvent être considérés comme des préconisations, sans avoir le statut de références validées.

Remerciements

Cet ouvrage a été réalisé sous la direction de Laurent Chevereau (Cerema - Direction technique Territoires et ville).

La rédaction principale de l'ouvrage a été réalisée par David Caubel et David Dubois (Cerema - Direction territoriale Centre-Est), sur la base des contributions de Nicolas Jouve (Cerema - Direction territoriale Nord-Picardie) et de Fabrice Reclus (Cerema - Direction territoriale Centre-Est).

Ont également contribué aux travaux préliminaires ayant permis la réalisation de cet ouvrage : Nicolas Jouve, Thomas Durlin et Samuel Verstraete (Cerema - Direction territoriale Nord-Picardie), Patrick Gendre (Cerema - Direction territoriale Méditerranée), David Caubel et David Dubois (Cerema - Direction territoriale Centre-Est).

Le Cerema remercie l'ensemble des relecteurs : Jean-Yves Coudray (Conseil Général de la Mayenne), Grégory Blanc-Bernard (Communauté Urbaine de Lyon), Sophie Tendeiro (Conseil Régional de PACA), Jean Seng (AFIMB), Sandrine Mathon (Toulouse Métropole), Jean-Marie Bourgogne (Ville de Montpellier), Fabrice Thiebaut et Laurent Bigou (Cerema - Direction technique Territoires et ville).

Préface

Open Data : deux jolis mots accolés, mélange de racines anglo-saxonne et latine, porteuses de deux concepts : l'ouverture, synonyme d'accessibilité, de liberté au sens de compétition, mais aussi de coopération et de création de « commons » ; et « data » un pluriel latin que l'on érige en catégorie singulière « la data », signifiant *le don* dans son sens d'origine et qui, numérisée, devient un nouveau pétrole pour les uns, un fluide plus complexe qu'il n'y paraît pour les autres.

Ce qui est en jeu c'est l'accès aux données ouvertes à la réutilisation; c'est-à-dire, au-delà de l'accès à l'information qu'elles constituent, leur usage pour constituer des bases de connaissances et réaliser des services utiles aux consommateurs, aux collectivités, aux marketeurs, aux innovateurs, et aux contributeurs citoyens de toutes sortes.

La démarche de l'ouverture questionne les détenteurs de données : *que vais-je perdre à les ouvrir ?* Puis si ces données sont sources de valeur : *pourquoi ne pas aller la chercher moi-même ?* Et puis, face aux poids lourds anglo-saxons devenus maîtres dans l'art de la captation et de l'usage des données pour se placer en intermédiaire prélevant une part croissante de la valeur : *suis-je bien armé pour cela ?* Enfin va-t-il falloir transformer le fonctionnement de ma structure sous la pression du partage et du libre accès aux données que je produis.

La démarche stimule les innovateurs et les acteurs de la multitude, trouvant là matière à création, à coopération, et observant la naissance d'un monde nouveau et ouvert, dans lequel les utilisateurs reprennent une part du contrôle du processus de définition des besoins, de conception, de contribution à la mise au point, de contrôle de la qualité perçue, et de l'amélioration des produits.

La démarche attire le monde politique et les collectivités ; une fois dépassées la peur de la perte de la maîtrise de l'information et la crainte de perte de vie privée pour leurs mandants, ils comprennent, enthousiastes ou réticents, tout l'impact potentiel sur l'organisation de la démocratie et le bon fonctionnement collectif.

La vérité est que les données et tous les mécanismes associés de traitement algorithmique (big data), de délocalisation (cloud), de sécurisation et de gestion de leurs usages, constituent une forme de nouveau réseau métabolique de l'économie et de la société, irriguant toute l'organisation de l'entreprise, des liens sociaux, et de la vie personnelle.

Leur ouverture est le levier pour entrer dans cette nouvelle ère du numérique associant « internet of everything » : les producteurs multiples de données, la cybersphère du big data et du cloud, et la gouvernance numérique, associant l'homme augmenté et la décision algorithmique.

Cette ouverture va se faire bien au-delà de l'ouverture des données administratives, pour aller vers une prise en compte de l'intérêt général comme moteur d'ouverture de données administratives, mais aussi de données publiques et privées dans des secteurs comme le transport, la santé, le logement, etc. Cela conduira à la constitution d'un espace mixte de données, privées ou appartenant aux « commons », avec des effets d'externalité sur l'espace des services ainsi que sur l'organisation et le fonctionnement de l'entreprise et des administrations.

Pour réaliser cette métamorphose de la donnée, le chemin est long et difficile, du fait des réticences compréhensibles des uns et des enthousiasmes audacieux et pressés des autres. Il apparaît pour cela nécessaire de se donner pour idéal un cap volontaire et ambitieux d'ouverture croissante des données, tout en prenant en compte son nécessaire encadrement résultant de la même notion d'intérêt public.

Il convient donc de se donner les moyens de progresser de façon sûre et durable dans la définition des jeux de données concernés, dans la standardisation des descriptions et des modalités d'accès, mais aussi dans la définition des droits d'usages et de la loyauté des services réalisés avec les données, et encore dans une précision progressive du droit des données.

Le travail présenté par l'équipe du Cerema s'appuie sur des entretiens réalisés avec des collectivités sur leur perception des enjeux et des freins, en se focalisant sur les données de mobilité, particulièrement porteuses de valeur d'usage. Le rapport présente un travail important de catégorisation des jeux de données selon leur nature, leurs usages et les formats utilisables, qui a contribué notablement aux travaux de réflexion au niveau national sur l'ouverture des données de transport. En sus, il renseigne sur les difficultés pratiques rencontrées, ainsi que sur les craintes fantasmées ou réelles sur l'ouverture des données. Enfin, il témoigne d'une approche prospective sur la trajectoire d'évolution de la société liée à l'ouverture des données et notamment sur l'émergence de nouvelles gouvernances qu'elle va susciter.

Le travail présenté – informatif, synthétique et didactique – est fort utile pour une appropriation progressive de la problématique de l'ouverture des données et de sa contribution à la démocratie, à l'économie, et à la qualité de vie.

Francis Jutand,
Directeur Scientifique de l'Institut Mines-Télécom
en charge de la recherche et de l'innovation.

Sommaire

Sommaire	5
Introduction	6
1 - L'Open Data au sein des collectivités françaises	7
2 - Vers l'émergence d'une nouvelle gouvernance locale et d'un modèle économique	12
3 - Comment surmonter les difficultés rencontrées ?	20
4 - Les clés du montage d'une plate-forme Open Data	36
5 - Mobilité : le choix des jeux de données	44
6 - Recommandations pour la mise en œuvre d'une plate-forme Open Data	58
Conclusion	61
Annexes	63
Annexe A - La grille d'enquête	64
Annexe B - Liste des collectivités et entreprises interviewées	68
Annexe C - La licence ODbL et la licence ouverte de Etalab	71
Annexe D - Bibliographie	73
Annexe E - Glossaire	77
Table des matières	78

Introduction

En 2014, une cinquantaine de collectivités françaises avaient déjà lancé une démarche Open Data visant à ouvrir les données publiques dont elles disposent. Initialement motivé par un souci de transparence, ce mouvement Open Data a évolué. Il vise aujourd'hui davantage des services innovants, permis par la réutilisation des données.

Face à ce phénomène, les collectivités s'interrogent. Le Cerema souhaite à ce titre contribuer à une meilleure connaissance des enjeux et des difficultés qu'elles rencontrent lors de la mise en œuvre d'une démarche Open Data. Au-delà de cette approche globale, la présente étude propose également un focus sur les problématiques de la mobilité. En effet, cette thématique est souvent mise en avant, en raison de la disponibilité des données, de la possibilité de créer des services innovants et de l'intérêt porté par les usagers sur ce qui fait partie de leur quotidien.

L'étude s'organise autour des points suivants :

Quels sont les enjeux de l'ouverture des données ? L'Open Data est-il réellement, comme on l'entend parfois, l'outil vertueux pouvant conduire à l'émergence d'une nouvelle gouvernance locale et de nouveaux modèles économiques locaux ou nationaux ?

Comment surmonter les difficultés que représente l'ouverture des données par les collectivités ?

Quelles sont les clés du succès pour mettre en œuvre une plate-forme de données ouvertes ?

Quelles sont les principales données relatives aux transports et à la mobilité mises à disposition par les collectivités ?

Comment bien appréhender les données de mobilité (type de données, contenu, formats...) ?

La finalité de ce rapport est de rassembler l'ensemble des éléments permettant de formuler des propositions et des recommandations sur les démarches Open Data.

Pour ce faire, le principal matériau est une série d'entretiens réalisés auprès d'acteurs de l'Open Data.

Ces entretiens ont porté sur l'émergence du projet, sur les freins qui peuvent être présents (lors du montage, du développement et de la pérennisation de la démarche), sur les résultats apportés par la plate-forme Open Data et enfin sur les perspectives d'évolution envisagées. Le détail du questionnaire est fourni en Annexe A - page 64.

Ces entretiens ont été réalisés en 2013 par le Cerema, notamment auprès de neuf collectivités locales ayant déjà mis en œuvre une plate-forme Open Data. Celles-ci représentent les différents niveaux institutionnels mettant en œuvre des plates-formes Open Data : la région, le département, les établissements publics de coopération intercommunale (EPCI) et les communes. Le détail des acteurs interrogés et des plates-formes analysées dans le cadre de cette étude est fourni en Annexe B - page 68.

D'autres matériaux ont également été mobilisés, comme le rapport sur les données routières réalisé par la direction territoriale Centre-Est du Cerema, une note de synthèse bibliographique réalisée par la direction territoriale Nord-Picardie du Cerema, la revue des plates-formes Internet Open Data, ainsi que la revue de la littérature sur un sujet en pleine effervescence. Les sources bibliographiques sont détaillées en Annexe D - page 73.

1 - L'Open Data au sein des collectivités françaises

1.1 - Rappel des définitions

Au sens propre du terme, « Open Data » signifie « données ouvertes ». La « donnée » est définie par Sylvie Servigne¹ comme « *l'information brute recueillie par un capteur ou générée par l'utilisateur lui-même* ». Certains considèrent que tout le monde peut produire de la donnée et que « *Aujourd'hui, 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 technologies. D'autre part, les exigences réglementaires poussent les acteurs publics à se doter de bases de données solides leur permettant d'éclairer leurs décisions.

Mais dans l'acception générale et sur les plates-formes Open Data existantes, les « data » recouvrent plus largement l'ensemble des « documents », qu'il s'agisse de données (brutes), d'informations (données interprétées ou résultant de traitements) ou encore de fichiers contenant du texte.

Les « données publiques » peuvent être définies comme les données détenues par les organismes publics et ne relevant pas de la vie privée ou de la sécurité nationale [Lecko, 2012]. En termes de droit, les termes d'« information » ou de « document » sont cependant préférés au terme « données », car moins ambigus. De même, le terme « public » peut être ambigu et compris soit comme la source d'une donnée (qui est un organisme public), soit comme la destination de la donnée (qui a vocation à être le grand public).

En France, la loi n° 78-753 du 17 juillet 1978 définit les « documents administratifs » comme « *les documents produits ou reçus, dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission* ». Les données relatives au secret-défense, au secret bancaire ou statistique, les données à caractère personnel sont exclues du champ des données publiques.

Le terme « Open Data » fait aujourd'hui référence à l'ouverture et au partage des données publiques, qui consistent à mettre à disposition de tous les citoyens, sur Internet, toutes les données publiques brutes qui ont vocation à être librement accessibles et gratuitement réutilisables [Etalab, 2013 a et b].

On parle aujourd'hui principalement d'ouverture des données publiques, 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 autorités publiques. Ainsi, cette étude ne concerne que les démarches Open Data portées par les collectivités locales françaises. 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 Maître de conférence au Laboratoire d'informatique en images et systèmes d'information (LIRIS – UMR 5205 CNRS / INSA de Lyon / Université Claude-Bernard Lyon 1 / Université Lumière Lyon 2 / École centrale de Lyon).

1.2 - Rappel de l'historique législatif

Dans les années 1960, la volonté citoyenne d'accéder à la donnée répondait au besoin de s'informer, de comprendre voire de contester l'action publique [*Décider ensemble*, 2012]. Elle a donné naissance, au début des années 2000, au mouvement Open Data, mieux organisé et aux objectifs plus larges, qui répond aux évolutions sociétales, notamment à la multiplication des données et aux nouvelles perspectives offertes par le numérique.

Ce mouvement, en plein essor depuis le milieu des années 2000, a trouvé un 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.

En France, la Déclaration des droits de l'homme et du citoyen de 1789 prévoit dans son article 15 que « *la société a le droit de demander compte à tout agent public de son administration* ».

En 1978, l'article 1^{er} de la loi n° 78-753 du 17 juillet 1978 fonde le principe de la liberté d'accès aux documents administratifs. Cette loi offre la possibilité aux citoyens de demander l'accès à ces documents et leur octroie la possibilité de se retourner, en cas de refus, vers la Commission d'accès aux documents administratifs (CADA). Certains documents, relatifs au secret-défense ou ceux contenant des données à caractère personnel, ne sont pas communicables. Dans son article 2, cette loi (dite loi CADA) précise qu'elle concerne également « *les établissements publics ou organismes, fussent-ils de droit privé, chargés de la gestion d'un service public* ».

→ **Concernant les transports**, les données récoltées dans le cadre d'une délégation de service public peuvent donc être demandées, sous réserve qu'elles ne portent pas « *atteinte au secret en matière commerciale et industrielle.* » (article 6). En l'absence d'une liste claire des types de données concernés, les avis formulés par la CADA apportent des éclairages.

Dans son avis n° 20083234, la CADA a ainsi considéré que les données de fréquentation annuelle d'une station de métro parisien « *relèvent des informations relatives aux conditions d'exploitation du service public assuré par la régie* », et que la RATP doit donc fournir ces données à qui le lui demande. La commission ajoute que même les informations « *susceptibles de révéler en partie les stratégies commerciales* » doivent être communiquées « *à toute personne qui en ferait la demande* » lorsqu'elles « *ont trait à la qualité et aux conditions d'exécution du service public concerné* »².

Par ailleurs, dans son avis n° 20100700, la CADA s'est prononcée sur le cas d'une base de données mise à disposition d'une administration publique par un tiers dans le cadre d'un abonnement payant. « *La commission considère que la copie d'écran du logiciel où figurent ces données peut être regardée comme un document administratif communicable* », et que les règles associées aux conditions d'utilisation de cette base « *ne font pas obstacle au droit d'accès ouvert par la loi du 17 juillet 1978* », puisqu'elles n'ont pas de valeur plus forte qu'une loi³.

Par la suite, différents textes de loi sont venus renforcer les obligations faites aux personnes publiques de rendre leurs données accessibles.

En 1998, trente-neuf États signent la convention d'Aarhus. Cet accord international vise à :

- améliorer l'information environnementale délivrée par les autorités publiques, vis-à-vis des principales données environnementales ;
- favoriser la participation du public à la prise de décisions ayant des incidences sur l'environnement (par exemple, sous la forme d'enquêtes publiques) ;
- étendre les conditions d'accès à la justice en matière de législation environnementale et d'accès à l'information.

2 L'avis complet est disponible sur le site de la CADA : <http://www.cada.fr/avis-20083234.20083234.html>

3 L'avis complet est disponible sur le site de la CADA : <http://www.cada.fr/avis-20100700.20100700.html0100700.html>

Ratifiée en France en 2002, cette convention a été déclinée dans la Charte de l'Environnement, adossée à la Constitution, ainsi que dans la directive européenne 2003/4/CE. Le portail <http://www.toutsurlenvironnement.fr/> est la réponse française au premier pilier de la convention.

→ **Concernant les transports**, ce portail informe notamment sur les impacts environnementaux des différents modes de transport.

En 2003, la directive européenne 2003/98/CE (dite directive PSI) 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 [Conseil National du Numérique, 2012]. Ce droit à la réutilisation des données publiques est transposé en droit français 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 du 6 juin 2005 rend le droit à la réutilisation opposable, la CADA étant chargée de veiller à son application.

→ Ce droit à la réutilisation ne s'applique néanmoins pas aux données obtenues dans le cadre d'une mission de service public à caractère industriel et commercial (SPIC), comme une Délégation de Service Public (DSP) **dans le domaine des transports**.

En 2007, la directive européenne 2007/2/CE, dite directive Inspire, vise à établir une infrastructure d'information géographique dans la Communauté européenne pour favoriser la protection de l'environnement. La directive s'adresse aux autorités publiques : l'État, les collectivités territoriales et leurs groupements, les établissements publics ainsi que « *toute personne physique ou morale fournissant des services publics en rapport avec l'environnement* ». Elle leur impose de rendre accessibles au public et aux autres membres de l'Union européenne – en les publiant sur Internet – les données géographiques qu'elles détiennent, dès lors que ces données sont sous forme électronique et qu'elles concernent l'un des 34 thèmes listés par la directive. La directive Inspire est transposée en droit français par l'ordonnance n° 2010-1232 du 21 octobre 2010.

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. La mission Etalab est aujourd'hui intégrée au sein du Secrétariat général pour la modernisation de l'action publique (SGMAP) créé fin 2012 par le gouvernement et placé sous l'autorité du Premier ministre. 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.

Ces dernières décennies, 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.

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 directive européenne 2013/37/UE modifiant la directive 2003/98/CE concernant la réutilisation des informations du secteur public.

Cette nouvelle 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 souhaitent rendre réutilisables ou non. Le périmètre d'application de ces principes couvre désormais toutes les données publiques et non 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 les outils informatiques. 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 au sens de la loi⁴ [Ecoter, 2013].

→ **Dans le domaine des transports**, la transposition de cette directive sera plutôt neutre.

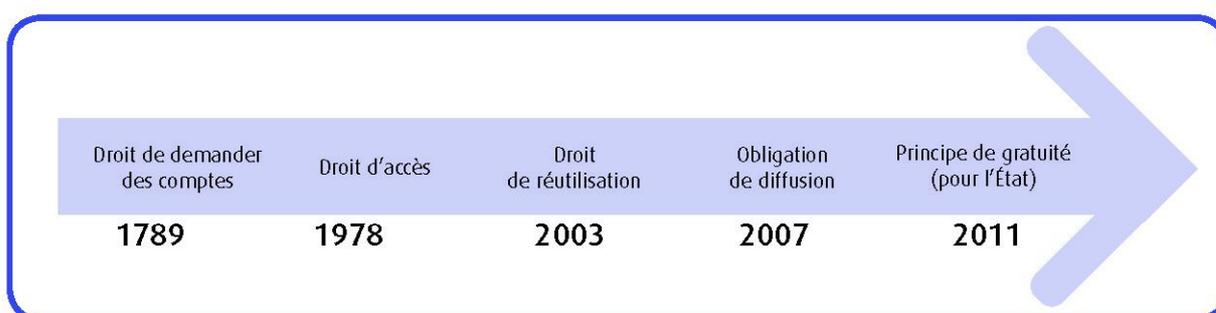


Illustration 1 : Principales évolutions législatives sur l'ouverture des données

1.3 - Où en est-on aujourd'hui en France ?

Le mouvement Open Data a pris de l'ampleur à la fin des années 2000. Aux États-Unis, la plateforme gouvernementale de partage des données www.data.gov est lancée en 2009 par le président Obama. La plateforme www.data.gov.uk suit en 2010 au Royaume-Uni, faisant écho à différentes initiatives de la société civile.

En France, les collectivités locales sont les premières à créer leur plateforme. En avril 2010, l'agglomération de Rennes met en ligne un catalogue de données ouvertes, *Rennes Métropole en accès libre*. D'autres villes ou agglomérations (Paris, Nantes, Bordeaux, Montpellier, Toulouse...) et plusieurs départements (Saône-et-Loire, Loire-Atlantique, Gironde...) lancent leurs plateformes courant 2011. L'État lance le portail interministériel www.data.gouv.fr fin 2011.

Début 2014, l'association LibertTIC répertoriait une cinquantaine d'initiatives d'ouverture de données publiques de la part des collectivités, avec des projets déjà opérationnels ou en cours de mise en place. L'annuaire Passim⁵ permet une connaissance précise des plateformes déjà mises

4 Article 10 de l'ordonnance n° 2005-650 du 6 juin 2005

5 Liste et détail des plateformes Open Data avec données de mobilité : <http://petitpois.passim.info/poi/search?k=open-data>

en service par les collectivités : une petite dizaine de Régions⁶ et de départements⁷ et une trentaine de villes et agglomérations de toute taille⁸.

Concernant les transports, on recense parmi ces plates-formes les horaires de 17 réseaux de transport. En plus de ceux-ci, certains réseaux fournissent des données sur les arrêts ou sur les lignes de transport. La plupart des plates-formes fournissent également des données sur les autres modes de transport (route, vélos...). Le détail et l'analyse des données disponibles figurent au chapitre 5 du présent document.

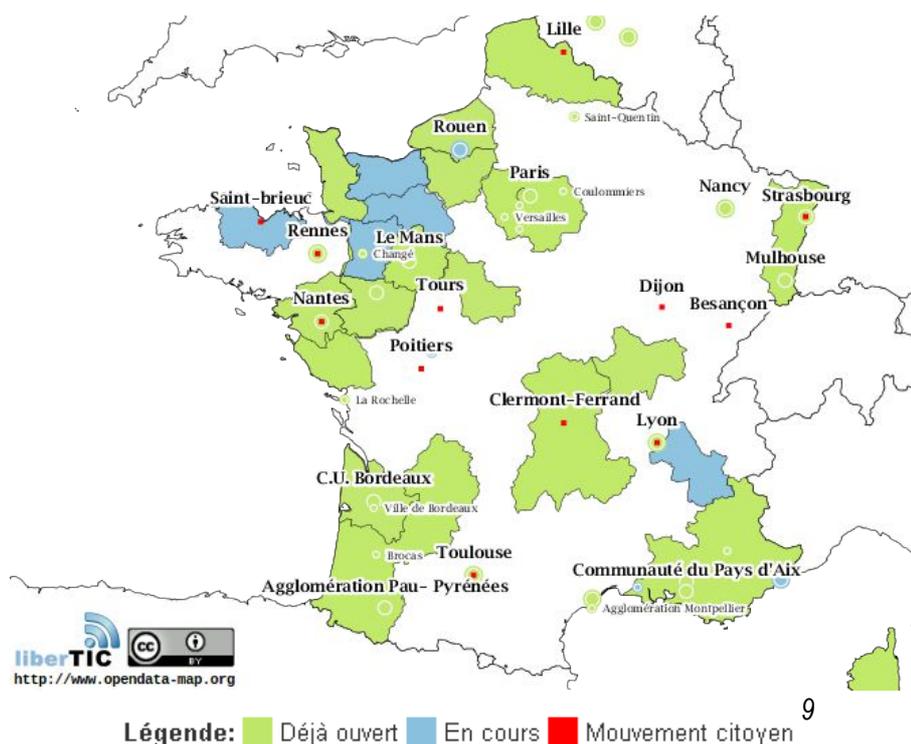


Illustration 2 : Cartographie des démarches d'ouverture des données en France
(source : LibeRTIC – www.opendata-map.org – mai 2014)

6 Corse, Alsace, Aquitaine, Auvergne, Île-de-France, Nord-Pas-de-Calais, Pays de la Loire et Provence-Alpes-Côte d'Azur
7 Eure, Gironde, Loir-et-Cher, Loire-Atlantique, Maine-et-Loire, Manche, Mayenne, Saône-et-Loire, Seine-Maritime et Hauts-de-Seine (le département de l'Oise ne dispose pas de plate-forme mais diffuse ses données sur la plate-forme nationale data.gouv.fr).

8 Certaines ont leur propre plate-forme (Angers, Bordeaux, Brocas, Changé, La Rochelle, Le Mans, Lyon, Montpellier, Mulhouse, Nancy, Nantes, Nice, Paris, Pau, Rennes, Strasbourg, Toulouse et Versailles), d'autres diffusent leurs données libres sur une plate-forme tierce (Aix-en-Provence, Coulommiers, Digne-les-Bains, Issy-les-Moulineaux, Longjumeau, Saint-Quentin, Sarlat-la-Canéda et Villecomble).

9 Mouvement citoyen : association ou collectif de citoyens œuvrant pour la promotion de l'ouverture des données

2 - Vers l'émergence d'une nouvelle gouvernance locale et d'un modèle économique

2.1 - Un enjeu politique et démocratique fort

« *Bien informés, les hommes sont des citoyens ; mal informés ils deviennent des sujets* ». Cette phrase d'Alfred Sauvy, démographe et sociologue français (1898 – 1990), pourrait résumer l'esprit de la démarche Open Data. L'Open Data contribue à donner la capacité aux citoyens d'être informés sur la vie publique, notamment par la lecture et la manipulation des données publiques, et donc de devenir acteurs.

L'un des objectifs des démarches Open Data les plus cités dans la bibliographie est celui de **transparence de l'action de l'État et des collectivités**. L'obligation pour la puissance publique de rendre compte à ses administrés est une valeur centrale des démocraties modernes. Cette responsabilité se concrétise notamment par un droit d'information des citoyens, et par une nécessité de transparence de l'État et des collectivités quant à leur fonctionnement [École des Ponts ParisTech, 2011]. 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 publiques¹⁰, publiée le 28 février 2013, parle 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)*. »

→ **Dans le domaine des transports**, la publication de données sur l'accidentologie¹¹ et sur la pollution¹² contribuent à cette transparence, mais reste aujourd'hui partielle. On ne recense pas de données sur le financement des projets de transports par exemple.

La mise à disposition de données brutes et leur réutilisation libre et ouverte permettent en outre d'**éclairer le débat public**. Un large accès aux données brutes confère par exemple aux journalistes une plus grande latitude pour la conduite d'investigations basées sur les faits et la construction d'argumentaires étayés (data journalisme). L'Open Data permet de recouper les données à titre de vérification, de les croiser avec d'autres sources pour mettre en perspective les phénomènes qui sont l'objet de leurs analyses [École des Ponts ParisTech, 2011].

10 Consultable à l'adresse www.gouvernement.fr/sites/default/files/fichiers_joints/donnees-publiques.pdf

11 Voir par exemple la plate-forme de Loire-Atlantique : <http://data.loire-atlantique.fr/donnees/detail/localisation-des-accidents-sur-les-routes-departementales-de-loire-atlantique/>

12 Voir par exemple la page Open Data de l'association Airaq, association de surveillance de la qualité de l'air en Aquitaine : <http://www.airaq.asso.fr/385-exportez-les-donnees.html>

Quelles données aujourd'hui ouvertes ? dans le domaine des transports publics (FR/UK/US)

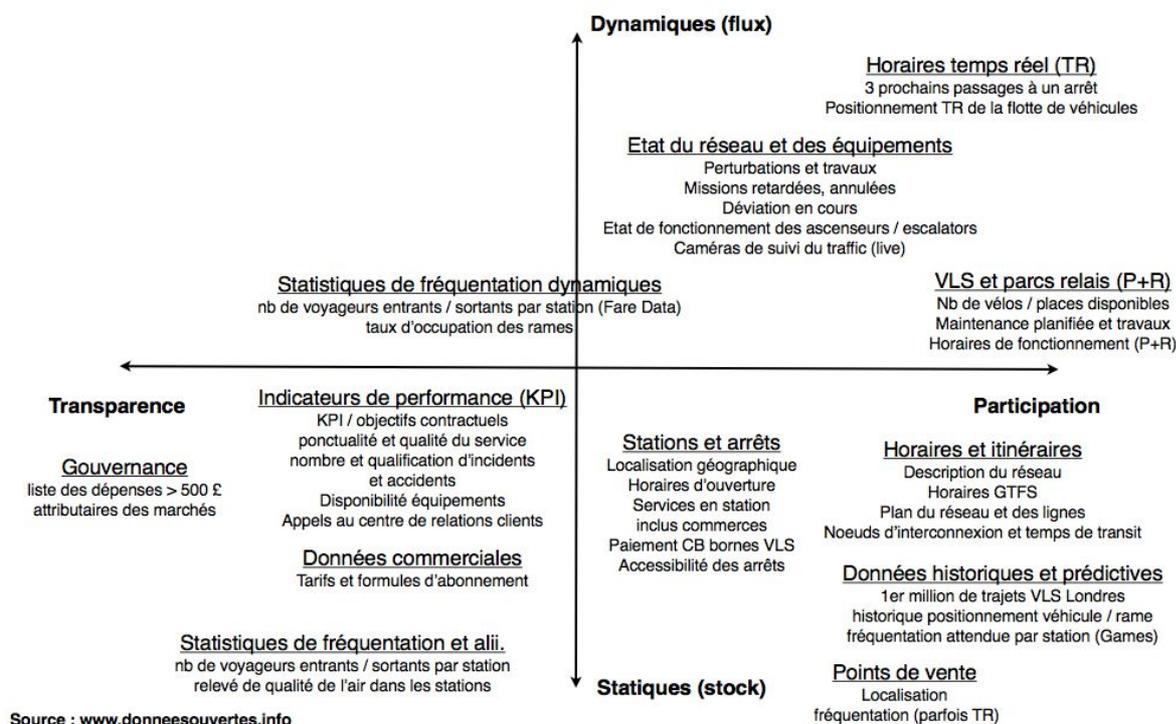


Illustration 3 : Caractérisation des données de mobilité selon leur contribution aux objectifs de transparence et de participation citoyenne (Source : www.donneesouvertes.info)

La mise à disposition de données brutes et leur réutilisation libre et ouverte permettent en outre d'**éclairer le débat public**. Un large accès aux données brutes confère par exemple aux journalistes une plus grande latitude pour la conduite d'investigations basées sur les faits et la construction d'argumentaires étayés (data journalisme). L'Open Data permet de recouper les données à titre de vérification, de les croiser avec d'autres sources pour mettre en perspective les phénomènes qui sont l'objet de leurs analyses [École des Ponts ParisTech, 2011].

Enfin, l'Open Data n'est pas uniquement une démarche descendante de la collectivité vers la société civile : il peut également aboutir à des phénomènes montants et **encourager l'implication des citoyens dans la vie publique**. La mise à disposition des données peut permettre l'émergence d'une réelle expertise citoyenne impliquée dans la vie publique [École des Ponts ParisTech, 2011]. La société civile peut alors s'emparer des données ouvertes pour créer des initiatives. En retour, la collectivité peut également bénéficier des compétences de la société civile et faire émerger ainsi des idées pour améliorer son fonctionnement ou créer de nouveaux services répondant aux besoins des citoyens [Décider ensemble, 2012].

2.2 - Une création de valeur économique pour les territoires

2.2.1 - L'économie de la donnée

La donnée publique constitue une matière première pour les industries de l'information et l'économie numérique [École des Ponts ParisTech, 2011]. Ainsi, la Commission européenne s'attache depuis plusieurs années à mesurer le marché des informations publiques [Chignard, 2012] et n'hésite pas à affirmer que « *les données publiques valent de l'or* » [Serdalab, 2012]. De nombreuses études concernant l'économie de la donnée publique ont déjà été publiées. À l'échelle de l'Union européenne, la valeur potentielle du seul marché de la donnée publique serait de 28 milliards d'euros en 2008 et les avantages économiques globaux liés à une plus grande ouverture des informations du secteur public représenteraient environ 40 milliards d'euros par an [Vickery, 2011]. Le potentiel de croissance qu'offrirait une plus large ouverture des données semble donc clair. Selon la Commission européenne, le total des gains économiques directs et indirects découlant de l'utilisation des données produites par le secteur public et des applications fondées sur ces données serait de l'ordre de 140 milliards d'euros par an [Commission européenne, 2011].

Certaines analyses, même si elles ne contestent pas l'existence d'un marché de la donnée publique, jugent toutefois ces estimations très optimistes [Serdalab, 2012 ; Chignard, 2012]. En outre, il n'est pas possible de relier l'économie de la donnée publique au sens large à l'économie de l'Open Data. En effet, les données libérées par les collectivités sur leurs plates-formes Open Data ne sont généralement pas celles qui font l'objet de la monétisation. La plupart des données « libérées » étaient déjà diffusées gratuitement sous différents formats, ce qui ne signifie pas qu'elles n'ont pas de valeur, mais que leur valeur est non monétaire [Chignard, 2012]. En tout état de cause, aucune étude n'a encore mesuré de gains économiques liés à l'ouverture de données publiques.

2.2.2 - Le marché de la donnée : des modèles économiques territoriaux encore fragiles

En parallèle aux enjeux démocratiques et d'action publique que représente l'accès aux données publiques, leur réutilisation par des développeurs, des entrepreneurs et des start-up est souvent présentée comme étant un enjeu économique, à court terme pour le développement du numérique, et à long terme pour la croissance et l'emploi [École des Ponts ParisTech, 2011]. La réutilisation des données publiques libérées permettrait de stimuler l'innovation dans les industries de l'information et l'économie numérique. En particulier, les données publiques pourraient devenir un ingrédient important d'encouragement au développement d'applications, entre autres mobiles.

Mais, malgré les promesses portées par le mouvement Open Data, les effets économiques et la création de valeur restent aujourd'hui difficiles à évaluer. Il est peu aisé pour les développeurs d'applications locales d'obtenir un retour sur investissement sans subventions de la part de la collectivité. En l'absence de modèle économique basé uniquement sur la réutilisation de la donnée, les développeurs utilisent principalement les plates-formes Open Data des collectivités locales pour se faire connaître : ils créent une application à partir d'une plate-forme Open Data en sachant pertinemment que les retombées économiques seront faibles, mais en espérant que cela leur permettra de faire reconnaître leur savoir-faire, et donc de décrocher d'autres marchés.

Une des pistes consisterait à dupliquer les mêmes services dans plusieurs territoires, ce qui suppose l'harmonisation des jeux de données des différentes collectivités.

Des applications pourraient ainsi voir le jour sur un territoire élargi, touchant un public plus large, et donc avec une meilleure rentabilité. Cette harmonisation peut se faire en utilisant des normes officielles validées par AFNOR ou des standards reconnus, comme le proposent la directive Inspire et la COVADIS¹³.

Il est toutefois réducteur de considérer les bénéfices économiques d'une démarche Open Data uniquement sous l'angle des applications mobiles développées à partir des données libérées [Chignard, 2012]. Ces applications constituent pour le grand public la partie visible de l'Open Data, mais il existe d'autres avantages, souvent indirects, notamment pour la collectivité elle-même.

→ Bien que particulièrement attendues, **les données relatives au transport** peinent également à trouver un vrai marché. L'utilisateur est en effet peu enclin à payer une information pour l'aider dans ses déplacements, considérant que celle-ci lui est due. Le modèle économique est donc difficile à trouver pour ces services, mêmes innovants. Néanmoins, des normes et des standards existent pour le trafic routier et pour le transport collectif: ils ne demandent qu'à être plus utilisés pour étendre le marché existant¹⁴.

2.3 - Des retombées positives pour la collectivité

L'émergence d'une économie de la donnée pour les territoires est controversée. En revanche, il est certain qu'une démarche Open Data ne peut pas être rentabilisée par la seule mise en place de redevance.

Dans tous les cas, la collectivité retirera de nombreux avantages d'une telle démarche.

2.3.1 - Mise à disposition des données : gratuité ou redevance ?

Pour la collectivité qui ouvre ses données, plusieurs modalités économiques sont envisageables concernant la mise à disposition et l'utilisation des données libérées.

La collectivité peut choisir d'instaurer une **redevance pour accéder à la donnée**. En effet, le droit français n'interdit pas systématiquement la facturation du coût de mise à disposition des données publiques. Cette autorisation de facturation du coût de la mise à disposition de la donnée, ainsi que celle de services à valeur ajoutée, est souvent importante pour les opérateurs dont la mission est de produire de l'information, et dont l'équilibre budgétaire peut dépendre de ces revenus complémentaires. Ce choix reste marginal parmi les collectivités interviewées dans le cadre de cette étude : seule une agglomération (Grand Lyon) rend potentiellement payant l'accès à une faible quantité de données¹⁵, les autres collectivités rencontrées mettant gratuitement à disposition l'ensemble des données libérées.

13 COVADIS : Commission de Validation des Données pour l'Information Spatialisée des ministères du Développement durable et de l'Agriculture.

14 Voir la fiche du Cerema « Normes et standards de l'information voyageur – des outils pour favoriser l'interopérabilité » (2014).

15 À la différence des autres collectivités interviewées, le Grand Lyon ne parle pas d' « Open Data » mais de « Smart Data ». Les données en accès payant représentent environ 1 % du total des données ouvertes. Cela ne concerne pour l'heure que les données de trafic routier en temps réel et l'historique du trafic heure par heure suivant 13 jours types.

La grande majorité des collectivités font le choix de la gratuité totale pour accéder à la donnée. Le cadre juridique et réglementaire, rappelé par le Premier ministre au cours du comité interministériel pour la modernisation de l'action publique du 18 décembre 2012 puis du Séminaire gouvernemental sur le numérique du 28 février 2013, prévoit en effet que la gratuité des données publiques est le principe par défaut. Pour certaines données, liées à l'obligation de rendre des comptes au citoyen, cette gratuité est un prérequis. Pour d'autres, l'expérience a montré que la mise à disposition de ces données gratuites favorisait la création de services à valeur ajoutée économique ou sociale, et donc l'émergence de nouveaux services au public et *in fine* un surcroît de revenus pour la collectivité.

Le 13 novembre 2012, le Groupement des autorités responsables de transports (GART) s'est prononcé en faveur de l'instauration d'une **redevance d'usage des données de transports publics**¹⁶. Cette prise de position pose la question de l'utilisation à des fins commerciales des données libérées gratuitement par les collectivités. En effet, le GART craint que certains acteurs – particulièrement les « *géants du web* » comme Google ou Apple – ne « *monétisent ces informations via l'intermédiaire de la publicité* ». Sans remettre en cause le principe de mise à disposition gratuite des données, le GART propose que les collectivités perçoivent de leur part une redevance calculée sur la base des revenus générés par la publicité. Cette prise de position du GART a suscité de nombreuses réactions sur la pertinence et la faisabilité d'une telle mesure.

En tout état de cause, il importe que d'éventuelles redevances d'accès ou d'utilisation des données ne créent pas de monopoles de fait ou des freins à l'innovation, notamment pour les jeunes entreprises. De fait, cela limite très fortement les avantages économiques directs que la collectivité peut en attendre. Parmi les collectivités interviewées, aucune ne cherche un modèle économique basé sur la monétarisation de la donnée ou de ses usages. Le choix du Grand Lyon de rendre potentiellement payant l'usage de certaines données ne vise pas des retombées économiques, mais a pour objectif de réguler le marché, en évitant la formation de monopoles.

2.3.2 - Des gains certains pour la collectivité : une modernisation des services publics

Les avantages monétaires directs qu'une collectivité peut attendre d'une démarche Open Data restent marginaux, voire nuls. Toutefois, l'Open Data apporte des bénéfices directs et indirects pour la collectivité.

Les collectivités réutilisent très fréquemment leurs propres données. Lorsque la donnée est accessible sur une plate-forme Open Data, il devient plus facile de l'utiliser en interne, car elle est référencée, décrite et à jour¹⁷. Cette possibilité d'**usage interne de la donnée** représente une source de réduction de coûts [Chignard, 2012] et facilite la transversalité entre les services. Le décloisonnement entre les services de la collectivité, et par extension avec les autres acteurs publics du territoire, est d'ailleurs souvent un apport attendu des démarches Open Data. 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)* ». L'ensemble des collectivités interviewées insiste sur les avantages procurés par la démarche Open Data en ce sens.

16 Proposition du GART accessible à l'adresse www.gart.org/S-informer/Salle-de-presse/Pour-une-redevance-liee-a-l-usage-des-donnees-des-transports-publics

17 Cet usage pourrait également être facilité par une plate-forme interne de données (non Open Data), néanmoins beaucoup de collectivités n'en disposent pas.

En interne à la collectivité, la mise en place d'une plate-forme Open Data nécessite un travail de capitalisation et de contrôle de la **qualité des données**, ce qui favorise la constitution de bases de données plus complètes, actualisées, mieux qualifiées et comportant des métadonnées. D'autre part, les services concernés doivent se réunir et échanger : le personnel prend ainsi connaissance des données à disposition dans les différents services mais aussi des différents modes de travail ou encore des difficultés rencontrées dans certains domaines [*Décider ensemble*, 2012].

Plus encore, la mise en place d'une plate-forme 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. Ces échanges peuvent aboutir à une amélioration de l'organisation interne de la structure publique¹⁸.

Par ailleurs, l'Open Data permet de **proposer de nouveaux services au public** en externalisant la réalisation de ces services à la communauté. En effet, la réutilisation des données par des développeurs, des entreprises ou des associations permet d'offrir aux citoyens de nouveaux services publics en ligne, souvent novateurs, que la collectivité n'aurait parfois pas su ou pu développer [École des Ponts ParisTech, 2011]. Cette externalisation peut déboucher sur une certaine économie financière pour la collectivité.

Enfin, l'ouverture des données est également profitable à la collectivité en termes **d'image et de communication**. L'Open Data reste encore « *une affaire de pionniers* » [Chignard, 2012], et apporte donc une image d'innovation et de transparence. Les éventuels dispositifs d'animation mis en place autour de la démarche Open Data contribuent également à une image positive. L'engagement de la collectivité en partenariat avec les acteurs de l'innovation, l'ouverture et la transparence participent à modernisation de l'action de la collectivité et de son rapport avec les usagers [École des Ponts ParisTech, 2011].

Rennes : Une image positive au service de l'attractivité du territoire

À Rennes, l'agglomération (Rennes Métropole) a changé son image grâce à sa plate-forme Open Data. Elle est maintenant considérée comme un territoire innovant. Rennes reste l'exemple emblématique de cet effet d'image, car l'agglomération était pionnière en France. Les retombées en terme d'image ne sont plus les mêmes pour les suivants. En général, cet effet d'image et la bonne connaissance de la plate-forme Open Data nécessitent une campagne de communication importante. Mais, dans le cas de Rennes, c'est la presse elle-même qui s'en est chargée. Ce côté pionnier de Rennes a ainsi intéressé revues et journaux, et l'agglomération estime avoir économisé environ un million d'euros sur deux ans en termes de publications (si Rennes avait dû payer les espaces de presse pour publier des articles à propos de sa plate-forme). L'Open Data et la communication faite autour du projet renforcent l'attractivité du territoire pour les partenaires privés.

¹⁸ Les plates-formes Prodigé (Plate-forme Régionale pour Organiser et Diffuser de l'Information Géographique) en sont un exemple.

2.4 - Les données de mobilité : un levier pour répondre aux exigences des usagers des transports

La mobilité est un secteur porteur de l'Open Data. La demande est forte : les données relatives aux transports et à la mobilité sont parmi les plus réutilisées et donnent lieu au développement du plus grand nombre d'applications mobiles [Chignard, 2012]. C'est notamment ce qui a été constaté lors des concours organisés à Rennes ou à Nantes afin de stimuler la réutilisation des données libérées.

Nantes, Rennes : Deux exemples contrastés de panel d'applications

Les applications pour smartphone référencées sur les plates-formes de Nantes et Rennes sont très différentes :

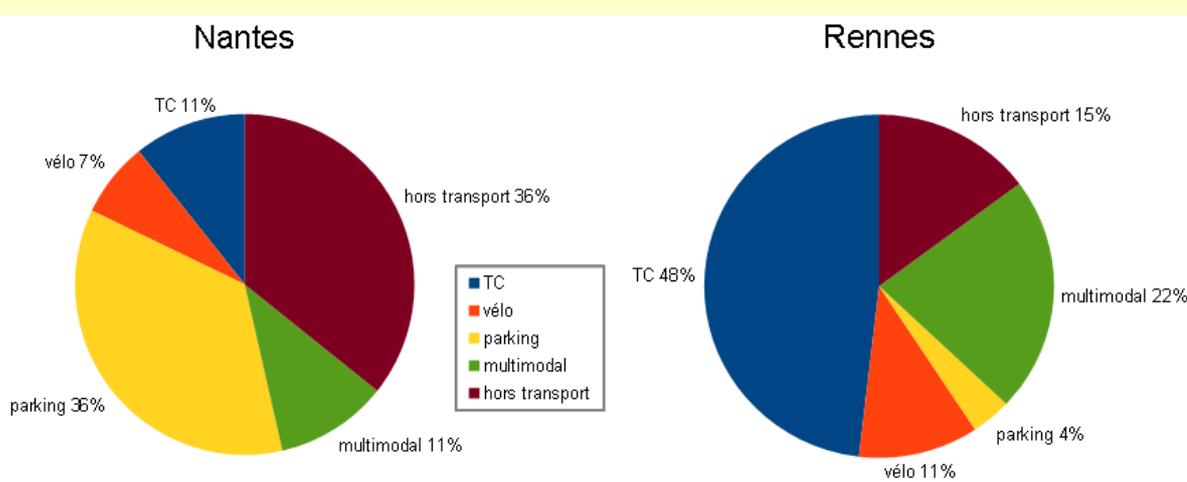


Illustration 4 : Thématiques des applications référencées sur les plates-formes Open Data (Cerema - mai 2014)

Ces différences s'expliquent en partie par les jeux de données mis à disposition. Ainsi, la mise à disposition de jeux de données sur les parkings publics de Nantes est à l'origine de nombreuses applications. En revanche, malgré des données semblables, la proportion des applications concernant les transports collectifs (TC) est bien plus faible à Nantes qu'à Rennes : à Rennes, on recense 19 applications basées sur les données d'horaires de transports collectifs, tandis qu'on n'en compte que 6 à Nantes.

Les données concernant la mobilité sont très variées : différents modes de déplacements (voiture, transports publics, vélo, marche...), différents types de données (descriptives, historiques, prédictives, en temps réel...), différents producteurs de données aux stratégies et aux intérêts parfois divergents (collectivités, opérateurs privés, usagers...). Elles sont détaillées au chapitre 5.

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.

En mettant en partage des données dont elle dispose déjà, la collectivité peut se concentrer sur la qualité de celles-ci, sans chercher à développer elle-même une application. L'expérience montre que bon nombre d'applications officielles sont moins performantes que des applications réutilisant des données ouvertes.

De plus, grâce à l'Open Data, des applications utilisant des données de thématiques différentes permettent l'émergence de services innovants, que le cloisonnement interne des collectivités n'avait jusque-là pas permis.

Certains mettent également en avant la possibilité pour une collectivité de faire corriger ses données gratuitement par les internautes eux-mêmes, en ouvrant des données de mauvaise qualité. Ceci peut être envisagé, mais nécessite de prendre des précautions et entraînera probablement des mécontentements.

Malgré tous ces avantages attendus, l'ouverture des données de mobilité ne va pas de soi et soulève de nombreuses questions. Selon la nature de la source des données – publique ou privée – des modalités d'ouverture différentes peuvent être envisagées. Le rôle des acteurs publics peut aussi en dépendre et s'orienter soit vers la fourniture de données, soit vers la fourniture de services. Un débat national sur l'ouverture des données publiques dans le domaine des transports a été lancé en avril 2014. Il a permis d'aborder la maîtrise des risques liés à l'ouverture des données transport, parmi lesquels le risque de création de situations monopolistiques. Le débat a également porté sur les types de données qui devraient être ouverts, les modes, et selon quelles modalités. Les recommandations issues de ce débat ont été remises en mars 2015 au Secrétaire d'État chargé des Transports, de la Mer et de la Pêche.

3 - Comment surmonter les difficultés rencontrées ?

Le terme de « freins » apparaît souvent dans les discours ou les écrits portant sur les démarches d'ouverture des données. Le tour d'horizon de la littérature et le retour des entretiens conduisent à préciser ce terme. Il s'agit plus globalement de difficultés qui sont plus ou moins aisées à surmonter, mais qui ne sont pas bloquantes. Bien souvent, les difficultés ne sont pas là où nous pensons les trouver.

Ceci amène à redire ce que de nombreuses collectivités, membres des réseaux de l'Open Data, affirment : « *Il n'y a pas forcément de freins, il suffit de faire ! Et de faire de façon pragmatique.* »

Les difficultés potentielles sont classées selon quatre catégories: les difficultés internes à la collectivité, celles d'ordre juridique, celles liées aux relations avec les partenaires, et celles plutôt techniques. Les collectivités ayant soulevé ces difficultés ou proposé des solutions pour les surmonter sont indiquées entre parenthèses.

3.1 - Des difficultés internes à la collectivité : convaincre de la finalité de l'Open Data

3.1.1 - De faux freins internes à la collectivité...

La première difficulté est liée directement au **manque de compréhension de ce qu'est l'Open Data** par les propriétaires des données (Rennes Métropole, Toulouse Métropole, Région PACA). Il s'agit de la seule véritable difficulté que peut rencontrer, en interne, une collectivité désireuse d'investir sur l'ouverture des données publiques.

Cette mauvaise compréhension de l'Open Data a des conséquences immédiates en termes de gouvernance interne de la collectivité. Les services, fournisseurs de données, peuvent être réservés quant à la motivation d'investir un domaine qui leur est inconnu (Grand Lyon). De plus, cela peut conduire à une méconnaissance ou un manque de lisibilité des données disponibles qui pourraient être publiées (Toulouse Métropole et Montpellier Agglomération).

Une fois cette compréhension commune établie, d'autres facteurs internes peuvent constituer un frein. Les collectivités mettent en évidence les trois suivants.

La première difficulté rencontrée est liée à l'appréhension des agents des collectivités animée par les conséquences de la **qualité de la donnée publiée**. Il apparaît une réticence des agents à mettre en ligne sur une plate-forme des données dont la qualité n'est pas optimale ou fiable à 100 %, ou qui ne sont pas mises à jour. Cette réticence exprime la peur d'être jugé par les utilisateurs. Publier une donnée reviendrait à « *exposer son travail à la critique* ». Cet argument est partagé par la majorité des collectivités rencontrées (Région Aquitaine, Ville et communauté urbaine de Bordeaux, Paris, Nantes Métropole, Ville de Montpellier, Toulouse Métropole).

La seconde difficulté, mentionnée également par une grande part des collectivités, est celle de l'expression par les agents d'une **Crainte d'avoir une charge de travail supplémentaire**.

Cette charge supplémentaire serait essentiellement liée à la construction de fiches de métadonnées décrivant les données ou à la mise à jour de l'information publiée, mais aussi à la gestion du « service après ouverture » visant à répondre aux questionnements des usagers sur la donnée. Ces tâches peuvent être perçues comme une charge supplémentaire, d'autant plus si les agents ressentent une pression hiérarchique ou politique pour répondre rapidement aux exigences ou demandes des utilisateurs.

Enfin, la troisième difficulté est une peur, de la part de l'agent, de **perdre sa maîtrise et son expertise de la donnée** : « *une perte de pouvoir* » (communauté urbaine de Bordeaux). En effet, en libérant une information, certains agents peuvent avoir l'impression de ne plus rester les seuls référents des données qu'ils maîtrisent. Et ce d'autant plus qu'ils ne savent pas ce que les utilisateurs feront des données publiées. Cela renvoie aussi à une préoccupation d'ordre psychologique de la part des personnes elles-mêmes et des services : celle de la responsabilité liée à l'usage qui est fait des données (Toulouse Métropole).

Cette crainte de perte de pouvoir apparaît également auprès des décideurs...

Extrait de la synthèse bibliographique du Cete Nord-Picardie¹⁹ (novembre 2013)

« *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 État "conservateur, adepte du secret, très centralisé" était également pointé du doigt par Gilles Babinet, ex-patron du Conseil national du numérique²⁰. Mais cette culture du secret ne peut caractériser l'ensemble des décideurs publics.* »

Pour autant, chacune de ces difficultés peut être débloquée plus ou moins aisément. Les collectivités rencontrées donnent des éléments de réponse pour remédier à ces difficultés.

3.1.2 - ... surmontables par le dialogue

Face à la méconnaissance ou l'incompréhension de ce qu'est l'Open Data, rien ne vaut l'**instauration du dialogue** et d'une **démarche explicative** pour convaincre les services internes des collectivités de l'intérêt de l'aventure de l'ouverture des données publiques.

À ce titre, le Grand Lyon insiste sur ce besoin d'expliquer la démarche de projet Open Data. Le discours doit permettre de montrer que d'autres sont engagés dans des démarches similaires, de montrer la plus-value, d'évaluer les risques et d'invoquer les usages possibles de l'Open Data en lien avec les politiques publiques.

De manière pragmatique, Nantes Métropole et Rennes Métropole se sont appuyées sur la communauté Open Data pour porter ce discours. Ainsi, les réticences internes à Nantes Métropole ont pu être levées grâce à l'intervention de la directrice de LiberTIC²¹ qui est venue expliquer lors d'un séminaire des cadres les enjeux de l'Open Data : « *Il faut l'intervention d'un professionnel externe, neutre et spécialiste du domaine, pour convaincre les agents.* ».

19 Depuis le 1^{er} janvier 2014, le Cete Nord-Picardie est devenu la direction territoriale Nord-Picardie du Cerema.

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

21 LiberTIC (www.libertic.wordpress.com) est une association nantaise de loi 1901, indépendante et non partisane, créée en décembre 2009 à la suite de l'*Open Data barcamp* de la Cantine à Paris. LiberTIC a pour objectif de promouvoir l'e-démocratie et l'ouverture des données publiques. Cette association développe ses activités sous les axes de la communication, la sensibilisation, la formation, l'animation et la production autour des données ouvertes.

Rennes Métropole a quant à elle organisé auprès de ses services des sessions *InfoLab*²² pour répondre aux interrogations suivantes : à quoi sert l'Open Data ? Quels intérêts ? Pourquoi ouvrir les données ? Quels usages ? Quelles sont les données disponibles ? *L'Infolab* tend à montrer que l'Open Data est un levier pour améliorer l'administration (gain en qualification sur la donnée, externalisation de services qui auraient dû être développés à un moment ou un autre, gain de productivité), pour travailler davantage en collaboration entre les services et avec les partenaires, mais aussi pour améliorer la connaissance des territoires.

Ce discours peut s'accompagner de la mise en place, quand les moyens le permettent, de référents Open Data dans les directions ou services internes (Montpellier Agglomération). Ces relais sont là pour lever les freins ou les difficultés liés aux différentes craintes évoquées précédemment.

Un des objectifs de l'Open Data est de jouer la transparence pour la collectivité. Ainsi le discours politique de Nantes Métropole a joué sur ce registre pour rassurer en affirmant qu'il n'y a pas à avoir peur de l'exposition à la critique, dans le sens où les remontées (critiques constructives) faites par les utilisateurs sont une occasion pour améliorer les données publiques de la collectivité. Ce n'est pas une fatalité, ni un jugement du travail des agents des collectivités. Au contraire, cela permet à la collectivité de faire croître l'expertise des agents et d'améliorer la qualité de ses données. En retour, elle pourra également fournir à l'utilisateur des données de meilleure qualité.

En complément des discours et des échanges avec les services internes des collectivités, un état des lieux peut être réalisé (comme l'a fait Toulouse Métropole), couplant les données que les services ont et pourraient ouvrir (l'offre) avec les besoins exprimés par les utilisateurs (la demande). On peut ainsi en déduire les données intéressantes à intégrer à la plate-forme Open Data.

Enfin, Toulouse Métropole rappelle que les problématiques de responsabilités (vis-à-vis de la qualité, de la fiabilité ou de l'usage fait des données) sont levées d'un point de vue juridique par la présence de licences²³ (ODbL, Licence Ouverte, autres...). Par ailleurs, un autre moyen de minimiser les craintes liées aux responsabilités de l'information diffusée est, comme le précise la Ville de Paris, d'anticiper les questions des usagers en décrivant, dès le départ, les données publiées (par des fiches de métadonnées) et en informant du degré de validité et de fiabilité des données. Ainsi, on minimise également la charge supplémentaire de travail qu'il adviendrait.

22 La Fing (Fondation internet nouvelle génération – <http://fing.org/>) organise des *InfoLab* visant à sensibiliser, faire découvrir ou lever les interrogations et doutes autour de l'Open Data.

23 La problématique du choix des licences est détaillée au paragraphe 3.2.2 page 24.

3.2 - Des difficultés juridiques : existence d'un cadre juridique

« *Il y a sûrement des peurs imaginaires. Mais il faut avancer et si un jour il y a de la jurisprudence, alors d'accord, on la prendra en considération. Il s'agit bien de prendre un minimum de risque pour innover.* » (Rennes Métropole). Cette citation incite à réfléchir sur la portée des difficultés juridiques qui peuvent se trouver là où on ne s'y attend pas.

La jurisprudence²⁴ est déjà fournie sur la protection des données à caractère individuel ou les droits d'auteurs. La question du choix des licences liées à l'ouverture des données peut être complexe mais ne pose pas forcément de vrai problème juridique, puisque les principales licences sont bien documentées et précisent bien les droits et devoirs des fournisseurs de données mais aussi des utilisateurs. En revanche, le cadre juridique n'est pas encore satisfaisant pour les collectivités qui souhaitent promouvoir, pérenniser ou soutenir des développeurs privés ayant mis en œuvre des applications à partir des données libérées.

3.2.1 - La protection juridique des bases de données

Les bases de données sont protégées par les droits de propriété intellectuelle. Ceux-ci comportent les droits d'auteur et les droits du producteur (droit *sui generis*).

Le régime du droit d'auteur protège toutes les œuvres de l'esprit, sous réserve que leur forme soit originale. Il s'applique dans de nombreux cas, sans qu'aucune formalité ne soit à remplir et quel qu'en soit le support matériel, et s'applique notamment à la structure des bases de données. Le droit du producteur des bases de données (droit *sui generis*) protège quant à lui le contenu d'une base de données, lorsque « *la constitution, la vérification ou la présentation de celle-ci attestent d'un investissement financier, matériel ou humain substantiel* »²⁵.

Les autorités publiques ne peuvent donc diffuser des données qu'à la condition qu'elles détiennent entièrement les droits de propriété intellectuelle. Ainsi, « *la loi de 1978, modifiée par l'ordonnance n° 2005-650 du 6 juin 2005, qui instaure le droit à la réutilisation des données publiques, exclut du champ de la réutilisation "les informations contenues dans des documents sur lesquels des tiers détiennent des droits de propriété intellectuelle".* » [Serdalab, 2012]

Il existe également une autre limite à la diffusion des données publiques : la protection des données à caractère personnel. En effet, les données détenues par les collectivités peuvent contenir des informations à caractère individuel ou personnel (nom de personnes, adresses, etc.). Ainsi, « *la loi du 6 janvier 1978 sur la protection des données à caractère personnel et la loi du 17 juillet 1978 imposent d'anonymiser ces données ou de réclamer le consentement des personnes nommées* » [Serdalab, 2012]. La donnée individuelle est régie par la loi et par la Commission nationale de l'informatique et des libertés (CNIL). Se pose alors comme difficulté en interne au service producteur et fournisseur de la donnée la question de la responsabilité et de la charge de l'anonymisation des données. Ce travail peut être fastidieux, mais il est obligatoire et soumis à des règles bien précises. Les collectivités rencontrées précisent bien leur préoccupation sur l'anonymisation des données : elle est, de fait, bien intégrée dans les plates-formes Open Data. Par exemple, Nantes Métropole explique que le jeu de données sur les prénoms les plus donnés ne reprend que les prénoms avec au moins 5 occurrences pour qu'on ne puisse pas retrouver facilement telle ou telle personne par recoupement.

24 Voir les décisions de la CADA (Commission d'Accès aux Documents Administratifs) : http://www.cada.fr/spip.php?page=recherche_avis.

25 Article L. 341-1 du Code de la propriété intellectuelle, codifié par la loi n° 98-536 du 1^{er} juillet 1998, transposant la directive européenne du 11 mars 1996.

3.2.2 - Le choix des licences

Le **choix d'une licence** encadre les pratiques de réutilisation des données libérées. Ce choix a un **impact sur le plan juridique** (droits et devoirs des fournisseurs et des usagers de la donnée) **et sur les modèles économiques liés à l'usage de l'information**. Le choix est donc aussi un choix politique porté par la collectivité. C'est d'ailleurs là que réside la principale difficulté que pointent les collectivités rencontrées.

Une démarche Open Data doit nécessairement s'accompagner d'une licence dite « libre », cédant au réutilisateur certains des droits afférents à la base de données.

Dans la plupart des collectivités rencontrées, il y a eu un débat autour du choix de la licence (voir Annexe C - page 71). Les collectivités font le choix parmi les licences libres existantes – les deux principales étant la licence ouverte de Etalab²⁶ et la licence ODbL²⁷ – ou alors optent pour la mise en place d'une licence « maison » créée par leur service juridique²⁸. La multiplication des licences possibles ne soustrait pas le débat autour du choix. Bien au contraire, puisque chacune de ces licences renvoie à un modèle économique et à une vision politique de l'Open Data.

En effet, comme le rapportent la Région Aquitaine, la communauté urbaine et la Ville de Bordeaux, le débat peut porter sur la conception politico-économique des licences :

- Les collectivités peuvent faire le choix d'une approche profondément déréglementée, dérégulée pour créer de l'innovation et faire baisser les coûts du marché. Cette vision libérale se rattache à la licence ouverte de Etalab (2011), une licence ouverte qui contraint le moins les usagers des données publiques. Elle est libérale puisque la seule contrainte est de mentionner la paternité et la date de la dernière mise à jour. Il est alors possible de reproduire, redistribuer, modifier et exploiter les données à titre commercial. Le Grand Lyon utilise ainsi la licence ouverte de Etalab en premier niveau (afin de pouvoir autoriser un usage commercial de certaines données), et des licences « maison » en second et troisième niveau d'usage des données. Depuis novembre 2011, la licence ouverte de Etalab s'applique à l'ensemble des réutilisations libres gratuites de données publiques issues des administrations de l'État et de ses établissements publics administratifs, à l'exclusion de tout autre licence.
- Les collectivités peuvent également faire le choix d'une approche plus collectiviste visant à développer et partager ensemble, notamment en mettant en commun les fruits de la réutilisation des données publiques. Il s'agit ici de la licence ODbL (licence qui permet de copier, modifier et de faire un usage commercial sous trois conditions : 1. citer la source ; 2. redistribuer sous des conditions de partage identiques les modifications ; 3. maintenir ouverte techniquement la base de données redistribuée, modifiée ou non).

Force est de constater que, malgré la difficulté de positionnement politique et économique autour des licences, une grande partie des collectivités utilise la licence ODbL, ou son adaptation au droit français. Le choix majoritaire de la communauté Open Data est donc celui d'aborder les démarches d'Open Data de manière collaborative et collectiviste, plutôt que libérale et ancrée sur le système des valeurs et du marché économique.

26 <http://www.etalab.gouv.fr/pages/licence-ouverte-open-licence-5899923.html>

27 Version originale (en anglais) : <http://opendatacommons.org/licenses/odbl/>

Traduction française, qui ne fait pas foi, mais peut servir de base pour élaborer une licence s'en inspirant : <http://www.vvlibri.org/fr/licence/odbl/10/fr/legalcode>

28 Si, initialement, il n'y avait que la licence ODbL (voir Annexe C - page 71), le débat a tout de même eu lieu au sein des premières collectivités à se lancer dans la démarche d'Open Data. Ainsi, Rennes Métropole avait préféré élaborer sa propre licence, puisque ODbL n'était pas encore adapté au droit français, avant de passer en 2013 à la licence ODbL.

Cette licence ODbL semble faire son retour, même si la licence ouverte de Etalab fut choisie par un bon nombre de collectivités (dont Bordeaux et Montpellier). Ceci étant, comme le précise l'association Regards Citoyens, les licences ouvertes et ODbL sont « *deux choix politico-économiques qui peuvent parfaitement cohabiter, assurant tous deux un véritable cadre Open Data aux ré-utilisateurs* ».

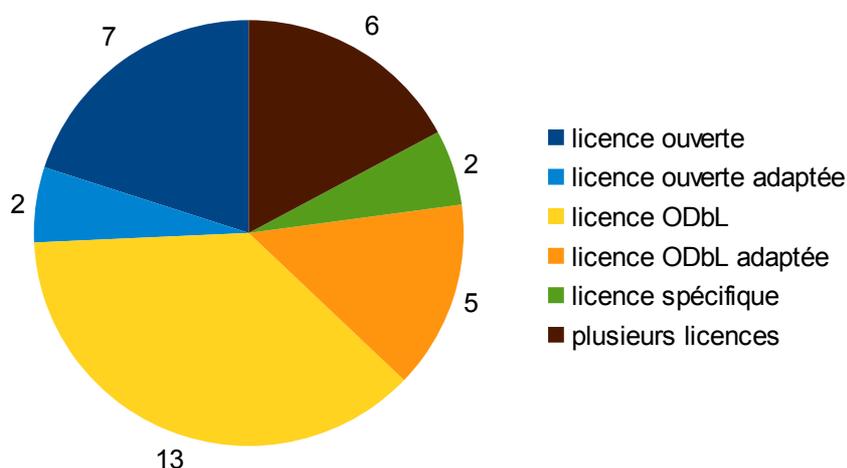


Illustration 5 : Répartition des licences utilisées parmi les 35 démarches Open Data ayant des données transports recensées en mai 2014 (Source : Cerema)

Dans un article paru en novembre 2013 intitulé « *Le fouillis des licences Open Data s'éclaircit* », la *Gazette des Communes* aborde cette question du choix des licences :

En deux ans, le paysage des licences compatibles avec un projet Open Data a bien évolué. On observe un resserrement sur deux licences libres, la licence ouverte d'Etalab et la licence ODbL. Le choix s'opère surtout en fonction de l'orientation politique et sociétale donnée à l'ouverture des données publiques.

Devant les nombreuses licences possibles, les collectivités souhaitant ouvrir leurs données peuvent se sentir un peu perdues. De l'une à l'autre, obligations et libertés accordées aux réutilisateurs ne sont pas les mêmes. Le choix de la licence doit se faire en fonction de l'objectif assigné à la politique Open Data mise en place. Étant entendu que, par définition, l'Open Data vise à faciliter la diffusion et la réutilisation des données publiques, sans faire de discrimination envers tel ou tel usager ou usage.

Pour cette raison, les licences dites libres sont le choix logique qui s'impose. Une licence libre octroie quatre droits : 1. celui d'utiliser, 2. d'étudier, 3. de modifier et 4. de redistribuer les modifications. [...]

Droit sui generis – *En principe, les contenus des bases de données mises à disposition en Open Data ne sont plus couverts par du droit d'auteur. La base de données elle-même peut être protégée par du droit d'auteur sous certaines conditions strictes sur l'originalité de sa forme. Une directive européenne de 1996 a rajouté un droit exclusif sui generis « pour protéger l'investissement, en temps, en moyens financiers et en travail, nécessaire pour établir une base de données, que celle-ci ait ou non un caractère intrinsèquement innovateur ».*

Pour l'heure, la situation n'est pas clarifiée, regrette Claire Gallon de LiberTIC, une association de promotion de l'Open Data : « Il y a eu un mouvement d'homogénéisation des licences autour de la licence ouverte (LO) et l'Open Database License (ODbL). On a ensuite vu une régression avec le développement des licences maison. L'intérêt des sigles ODbL et LO était de ne pas avoir à relire les conditions juridiques. Aujourd'hui, on trouve des licences appelées LO ou ODbL alors que le texte juridique est différent, c'est nuisible à toutes les démarches. » « Il y a eu stabilisation, mais il est vrai que des acteurs privés ont adapté ces deux licences », précise Simon Chignard, consultant sur l'Open Data. Benjamin Ooghe-Tabanou, de Regards citoyens, une association militant pour l'Open Data, penche aussi pour un resserrement autour de ce duopole.

« Aussi simples à appréhender et à accepter que possible » – *Dans son rapport sur les redevances sur les données publiques, le magistrat de la cour des comptes Trojette souligne l'importance de ne pas noyer les utilisateurs dans les méandres de licences spécifiques :*

« Dans un contexte de concurrence avec des données produites hors du service public, les contraintes ainsi imposées aux réutilisateurs doivent être aussi simples à appréhender et à accepter que possible : une jeune pousse du secteur des Big data, désireuse de réutiliser massivement les données publiques de vingt administrations différentes dans quinze pays différents et de les croiser avec des données privées ne consentira vraisemblablement pas à faire appel aux services d'un avocat pour étudier toutes les licences types non standards de l'administration. Dès lors, l'administration devrait privilégier le recours à des licences types avec lesquelles d'éventuelles licences spécifiques seraient automatiquement compatibles. Ces licences types pourraient avoir l'avantage d'une formulation compatible avec les standards internationaux – pour la définition desquels la France a un rôle à jouer – et faire l'objet d'une traduction pour accroître l'attractivité des plateformes hors de France. »

Tableau 1 : Extraits d'un article de la Gazette des Communes (source : La Gazette des Communes, Sabine Blanc, 25 novembre 2013)²⁹

²⁹ La suite de l'article (disponible au lien suivant <http://www.lagazettedescommunes.com/208893/le-fouilli-des-licences-open-data-seclaircit-fiche-pratique/>) présente les différentes licences : Etalab, Open Data Commons (dont ODbL, ODC-By, PDDL), Creative Commons (CCo), et les licences françaises qui sont tombées en désuétude (LIP).

Le choix des licences n'est donc pas une simple affaire pour les collectivités. Et même si les pratiques se resserrent autour de la licence OdbL et de la licence ouverte de Etalab, la difficulté rencontrée a été celle de **l'adaptation de la licence au droit français**. Cette difficulté a été surmontée par la communauté Open Data présente auprès des collectivités et par leurs services juridiques. Comme le souligne la Ville de Paris, « *on peut toujours aller au-delà [des difficultés] grâce à la communauté* ». Parmi les acteurs de la communauté Open Data les plus actifs, on peut citer :

- l'association *Open Knowledge Foundation* (OKFN - <http://fr.okfn.org/>), qui vise à promouvoir le partage de connaissances en Europe. Elle a développé trois types de licences orientées bases de données, basés sur le droit anglo-saxon (licences Open Data Commons) ;
- l'association *Veni Vidi Libri* (<http://www.vvlibri.org/>), qui a pour objectif de promouvoir les licences libres ainsi que de faciliter le passage de créations sous licence libre. C'est d'ailleurs cette association qui a transposé la licence ODbL au droit français ;
- l'organisation *Creative Commons* (CC)³⁰ ;
- l'association LiberTIC (<http://libertic.wordpress.com/>).

Les collectivités peuvent également s'appuyer sur l'association OpenDataFrance. Celle-ci « *a pour but de regrouper et soutenir les collectivités engagées activement dans une démarche d'ouverture des données publiques et de favoriser toutes les démarches entreprises par ces collectivités dans le but de la promotion de l'Open Data* »³¹.

Certaines plates-formes n'associent pas une licence unique à l'ensemble de leurs jeux de données. C'est le cas des plates-formes mutualisant les données de différents producteurs, qui peuvent avoir chacun une politique différente de diffusion de leurs données. C'est également le cas de plates-formes qui adaptent la licence au contenu même de la donnée.

Ainsi, le Grand Lyon a établi plusieurs niveaux de licences selon les données, et les usages potentiels de ces données par les utilisateurs ou clients, porteurs de démarche économique locale. Même si le Grand Lyon affirme que la multiplication des licences complique la réutilisation des données, il a fait le choix de mettre en place trois licences différenciées :

- la licence de premier niveau : la **Licence Ouverte** (Etalab). L'objectif est de favoriser les réutilisations commerciales des données. En juin 2014, 95 % des données sont mises à disposition, sans condition ni limitation, ni clause de partage à l'identique via cette licence ;
- la licence de second niveau : la **Licence Engagée**. L'objectif est de s'assurer que les réutilisations de quelques données stratégiques soient compatibles avec les politiques publiques et l'intérêt général. Les données sont mises à disposition gratuitement après authentification des utilisateurs ;
- enfin la licence de troisième niveau : la **Licence Associée**. L'objectif est de favoriser le développement de nouveaux services et entreprises émergentes ainsi qu'un écosystème concurrentiel équitable. Il est possible de mettre en place une redevance à partir d'un seuil élevé de volume d'activité du produit ou du service créé afin d'éviter les situations

30 *Creative Commons* est une organisation à but non lucratif qui facilite le partage et la réutilisation de la créativité et de la connaissance à travers des outils juridiques gratuits. Installée à San Francisco depuis 2001, elle a lancé la première version des licences CC en 2002. Le réseau affilié à *Creative Commons* est constitué de plus de 100 institutions travaillant dans plus de 70 juridictions pour développer les licences et les activités de CC dans le monde. Depuis 2003, l'institution affiliée à *Creative Commons* en France est le CERSA (Centre d'Études et de Recherches en Sciences Administrative et politique), unité mixte CNRS/Université de Paris II, qui anime notamment le site <http://www.creativecommons.fr/>

31 <http://opendatafrance.net/>

de monopole. L'absence de redevance est possible dans le cadre d'expérimentations, pour les universités, les acteurs de la société civile et les PME innovantes.

La mise en place de ces deux licences innovantes au Grand Lyon (Engagée et Associée) a été rendue possible grâce à ses propres services juridiques. Cette compétence juridique au sein du Grand Lyon est d'autant plus importante qu'elle est absente dans la plupart des communes de la communauté urbaine. Les communes sollicitent souvent un appui juridique du Grand Lyon, pour comprendre et adhérer au modèle économique.

→ La question d'une **licence spécifique aux données de transport** est parfois évoquée. Celle-ci pourrait exiger des réutilisations conformes à l'intérêt général et aux politiques publiques. Ceci pourrait apaiser les collectivités qui craignent un mauvais usage des données qu'elles ne pourraient pas contrôler. Elles craignent par exemple qu'une application pour smartphone propose un itinéraire de contournement routier passant devant une école, ou encore dévie des centaines d'usagers du RER vers une navette de 30 places qui passe une fois par heure.

3.2.3 - Un véritable frein : le code des marchés publics

Malgré les difficultés juridiques et économiques liées au choix des licences, le véritable obstacle juridique se trouve dans le Code des marchés publics. Plusieurs collectivités le mettent en évidence (Région Aquitaine, la communauté urbaine et la Ville de Bordeaux, Nantes Métropole et Rennes Métropole).

Les collectivités ne savent pas **comment aider les développeurs qui ont besoin de modèle économique viable**. Si, techniquement, les plates-formes mutualisées peuvent contribuer à l'émergence de modèles économiques (notamment par la réplique d'applications sur des territoires larges, et donc touchant plus de clients), la difficulté réside dans la pérennisation des services développés.

À ce titre, la forme actuelle du **code des marchés publics rend difficile, si ce n'est impossible, le soutien financier aux API et applications**. En effet, les besoins de développement et de pérennisation émergent a posteriori, c'est-à-dire une fois que l'application est réalisée. Par exemple, comme le précise Nantes Métropole, lorsque une application lauréate d'un concours intéresse une collectivité, cette dernière ne peut acquérir facilement cette application. Elle doit passer par la procédure des marchés publics (et donc par une mise en concurrence selon les seuils des marchés publics). Or la réponse au marché public à mettre en place est déjà présente et connue par la nature de l'application mise en service.

Le Code des marchés publics rend donc impossible, dans la limite des seuils de marchés qu'il impose, l'attribution a posteriori d'un marché à un développeur ayant déjà mis en place un service répondant aux besoins exprimés. Cela complexifie la relation avec les développeurs. La collectivité ne peut pas « récompenser » le développeur en achetant son application, ou en essayant de pérenniser les services développés par un tiers. Cela renvoie à une réflexion plus globale et nationale sur la pérennisation d'un modèle économique émergent. C'est ainsi que de nombreuses applications, même récompensées, ne sont pas pérennes.

3.3 - Des difficultés partenariales : les jeux d'acteurs

Les collectivités peuvent rencontrer des difficultés externes à leur structure dans la mise en œuvre d'une démarche Open Data. La collaboration entre collectivités est souvent connotée d'enjeux politiques, ce qui complique la mise en œuvre ou la mutualisation des informations sur une plate-forme Open Data. De plus, la relation entre la collectivité et les fournisseurs de données peut aussi être conflictuelle. La confrontation est souvent liée à la nature et au contenu de la contractualisation (convention, délégation de service public, nombre d'intermédiaires), mais aussi aux logiques économiques différentes.

Le **dialogue et la négociation sont les éléments incontournables** pour de bonnes relations partenariales dans une démarche Open Data.

3.3.1 - Mutualisation ou stratégie d'acteurs entre les collectivités potentiellement actrices d'une démarche commune d'Open Data ?

Les démarches de mutualisation entre collectivités peuvent apporter plus de services et des économies de moyens, mais leur mise en œuvre est délicate.

Le principal frein évoqué par certaines collectivités interviewées réside dans la **difficulté de collaborer et de rassembler un ensemble de collectivités locales autour d'une démarche commune** d'Open Data (Région Aquitaine, communauté urbaine et Ville de Bordeaux). Et ce même si c'est un objectif initialement affiché par les parties prenantes. Bien souvent, chaque collectivité veut avoir son propre portail et ne cherche ni une harmonisation entre les portails au sein d'un même supra-territoire, ni un référencement des données des autres collectivités partenaires. Cette velléité correspond à des objectifs politiques de visibilité et de communication pour se montrer à la pointe de l'actualité du mouvement (« *susciter le buzz* »).

Par ailleurs, bien souvent, la frilosité des partenaires d'une collectivité est le fruit du positionnement politique des élus. Ainsi, comme le met en avant le Grand Lyon, même si les services techniques de certaines communes ou collectivités sont très motivés pour adhérer à une démarche Open Data, ils sont bloqués et attendent un feu vert politique qui peut venir plus ou moins rapidement. Les périodes récurrentes des élections (et périodes de réserve pré-électorale) sont aussi la source de la non-expression des élus sur ces démarches, de peur de la sanction électorale. Enfin, la réticence peut venir du fait que les démarches Open Data ne sont pas nécessairement perçues comme étant une question prioritaire.

Il semble que cette délicate mutualisation peut tout de même trouver une issue favorable. Ce fut le cas pour la Région Pays de la Loire, le département de Loire-Atlantique, Nantes Métropole et la Ville de Nantes. Ces quatre collectivités ont chacune leur site Internet Open Data, mais sur une trame et charte graphique harmonisée, avec des liens (référencements) entre les sites Internet des quatre collectivités. Cela permet à la fois une mutualisation des informations et le maintien d'une visibilité auprès des internautes. La relation entre ces quatre partenaires a été facilitée par une convention très précise signée au lancement du projet et qui détaille :

- les droits et obligations de chacun, et notamment la participation financière ;
- les tâches à réaliser par chaque partenaire (afin de garantir une bonne répartition du travail et du temps à consacrer) ;
- les grands enjeux stratégiques qui doivent guider le développement de la plate-forme ;
- le choix de la licence ;
- les conditions à remplir pour un nouvel entrant (qui peut être un partenaire privé).

Collectivités locales	Type de collectivité	Nom de la plate-forme Open Data	Lien internet de la plate-forme Open Data
Ville de Nantes et Nantes Métropole	Commune et communauté urbaine	Nantes - ouverture des données	http://data.nantes.fr/accueil/
Département de Loire-Atlantique	Département	Loire-Atlantique - ouverture des données publiques	http://data.loire-atlantique.fr/accueil/
Région Pays de la Loire	Région	Région Pays de la Loire - ouverture des données publiques	http://data.paysdelaloire.fr/accueil/

Tableau 2 : Exemple de mutualisation des plates-formes Open Data en Pays de la Loire

Malgré l'exemple précédent, la construction de démarches communes entre différentes collectivités n'est pas la règle. La mutualisation de l'information et le référencement³² ne sont pas la règle non plus. Pourtant, un modèle économique gagnant-gagnant pourrait émerger : gagnant pour les collectivités elles-mêmes réalisant ainsi des économies d'échelles sur les coûts et les moyens à mettre en œuvre ; gagnant pour les utilisateurs des données, qui pourraient proposer des services en tenant compte des différentes échelles territoriales, et **toucher ainsi un public plus large**.

Un cercle vertueux permettrait alors à la collectivité de bénéficier en retour de ces services à plus large échelle créés par les utilisateurs : autant de services qu'elle peut offrir aux citoyens sans avoir besoin d'investir.

La Ville de Paris porte un discours incitant les collectivités à ouvrir leurs données, a minima en utilisant les sites et les plates-formes déjà existants pour diffuser celles-ci³³. Il s'agit d'inciter à la mutualisation et la collaboration entre les territoires afin d'obtenir des économies d'échelles à moindre coût.

Au niveau européen, la directive Inspire³⁴ impose ce type de fonctionnement aux États membres, avec un portail national d'une part et un portail fédérateur européen d'autre part.

3.3.2 - Des logiques économiques et politiques différenciées entre les collectivités et les fournisseurs de données extérieurs

Pour un producteur, la donnée peut être perçue comme une source de revenus directs, et une démarche Open Data comme une campagne publicitaire. Cette perception peut fortement diverger de la vision d'une collectivité. La mise à disposition de données produites par un tiers peut ainsi être mise à mal pour des raisons économiques et politiques, y compris dans le secteur privé.

Les fournisseurs peuvent être très intéressés par les données libérées des autres, mais ne veulent pas forcément mettre à disposition leurs propres données. Et ce d'autant plus qu'ils peuvent entrer dans une vision de création de valeur économique liée à l'usage de la donnée.

32 Le référencement consiste à informer de l'existence de données, mais en renvoyant l'internaute sur une autre plate-forme pour le téléchargement. Ceci évite les doublons sur un territoire, qui engendreraient une charge de travail double et un risque d'avoir des données différentes selon la source.

33 Soit au niveau local (avec la plate-forme régionale <http://data.iledefrance.fr/>), soit au niveau national (avec la plate-forme <http://www.data.gouv.fr/>).

34 La directive INSPIRE est la directive 2007/2/CE du Parlement européen établissant une infrastructure d'information géographique dans la Communauté européenne (voir : http://fr.wikipedia.org/wiki/Infrastructure_for_Spatial_Information_in_the_European_Community)

Cela signifie qu'ils ne sont pas forcément prêts à mettre à disposition leurs données pour des raisons d'incertitude sur ce qui peut en être fait et sur l'activité commerciale potentielle. La valeur marchande de la réutilisation des données échapperait à son producteur. Un mauvais usage des données pourrait aussi nuire à l'image du producteur et avoir des conséquences économiques négatives indirectes.

De plus, quitte à ouvrir leurs données, les producteurs préfèrent renvoyer les usagers sur leur propre site ou plate-forme de données, leur permettant d'être plus visibles, voire d'engranger des revenus par la publicité.

Enfin, c'est parfois le propre modèle économique des fournisseurs qui bloque la mise en partage des données (par exemple, le fonctionnement des établissements publics que sont l'IGN ou l'Insee a longtemps reposé sur des ressources financières assises sur l'accès payant à leurs données). Cela peut être aussi la politique générale vis-à-vis des concurrents ou d'une situation de monopole qui peut être un facteur bloquant (par exemple, la SNCF). Toutefois, les lignes évoluent dans les deux cas de figures : l'IGN évolue rapidement vers l'Open Data³⁵ ; l'Insee met davantage de fichiers « détail »³⁶ et de données à disposition sur son site internet ; et la SNCF a ouvert sa plate-forme Open Data en 2012³⁷.

3.3.3 - Le poids du lien contractuel entre les collectivités et les fournisseurs de données

La difficulté des relations entre la collectivité et un prestataire, notamment dans le domaine des transports et de la mobilité, réside essentiellement dans l'absence d'inscription de la mise à disposition des données dans les conventions partenariales ou les délégations de services publics (DSP) pour le transport urbain. Même si la collectivité est (co-)propriétaire des données avec le délégataire, il est relativement difficile pour elle du fait de cette absence, de demander l'ouverture des données transports. Le prestataire souhaite généralement rester propriétaire de la donnée, et ce même si les moyens techniques existent pour la récupérer et la mettre à disposition.

La plupart des collectivités interviewées sont confrontées à la difficulté liée au manque de précision du volet « données » dans les conventions, les contrats ou les DSP. Les prochains renouvellements de contrats, conventions ou DSP seront l'occasion d'y mentionner la récupération des données pour alimenter les portails Open Data. Mais, en attendant, les collectivités souhaitent déjà proposer ces données sur leurs portails. Elles entament donc des négociations avec les autorités organisatrices des transports urbains (AOTU), les exploitants et les prestataires de services de mobilité (vélo en libre-service, gestionnaires des parcs de stationnement, automobile en libre-service ou en partage...).

Ces négociations sont plus ou moins complexes selon le nombre d'acteurs intermédiaires entre la collectivité et le prestataire de service de mobilité. Par exemple, à Rennes Métropole, le fait de ne pas avoir d'intermédiaire avec les délégataires facilite la coordination et la mise à disposition des données transports (Keolis Rennes).

35 www.ign.fr/institut/lign-lopen-data

36 En mai 2014, l'Insee diffusait ainsi sur son site internet 71 fichiers « détail » : des bases de données électroniques comportant les enregistrements individuels, anonymisés, afférents à des enquêtes ou opérations statistiques réalisées par l'Insee.

37 www.data.sncf.com

À l'opposé, le Grand Lyon doit passer par l'AOTU, en l'occurrence le Sytral, pour négocier la mise à disposition des données de l'exploitant (Keolis Lyon). Les négociations dépendent également de la nature et de la gestion des contrats entre AOT et délégataires. Il y a souvent une peur de ne pas maîtriser l'information et ce qu'on en fait : le monde de l'information libre semble aller trop vite par rapport à la temporalité de l'action publique (problèmes d'images et de responsabilité).

Par ailleurs, les producteurs des données peuvent également être ouverts au partage de l'information, mais à la condition de maîtriser celles-ci. De nombreuses collectivités interviewées rapportent leur expérience avec leur opérateur de vélos en libre-service. La propriété partagée des données a rendu complexe l'ouverture des données. Les avancées des démarches Open Data ont ensuite séduit l'opérateur, qui propose aujourd'hui sa propre plate-forme Open Data³⁸. Malgré cela, peu de « passerelles » existent entre cette plate-forme et celles des collectivités. Au mieux, on peut trouver un référencement des données. De plus, les licences ne sont pas nécessairement identiques entre l'opérateur et les collectivités, ce qui complique la compréhension et l'usage des données.

3.4 - Les difficultés techniques : tout tourne autour de la donnée

Les utilisateurs des plates-formes Open Data disposent généralement d'un espace d'expression. On y retrouve leurs principales préoccupations : l'accès aux données des plates-formes, la compréhension et l'interprétation des données, ou encore la réutilisation de ces données. Il s'agit là de véritables freins au développement de services réutilisant les données ouvertes. Ces freins ne doivent pas constituer un prétexte pour ne pas ouvrir ces données. En effet, les progrès techniques ainsi que les réflexions de la communauté Open Data sont porteurs de solutions et font émerger les préconisations nécessaires pour les futures démarches d'Open Data. En revanche, la collectivité doit mettre l'accent sur la pédagogie autour des données et communiquer sur les outils permettant de les manipuler.

3.4.1 - Le choix des formats, frein à l'usage des données ?

La première question pour la diffusion d'un jeu de données réside dans le choix des formats. Si, initialement, il n'y avait pas beaucoup de recul sur la question, aujourd'hui, on observe une très grande diversité des formats des données. Cela entraîne nécessairement d'autres difficultés en aval : la plus ou moins grande difficulté d'usage des données, leur poids, les compatibilités, l'appariement de différents jeux de données, etc. Comme le précise Montpellier Agglomération, certains formats sont difficiles à exploiter pour les non-spécialistes, alors que les démarches d'Open Data s'adressent globalement à tous les citoyens.

Le choix du format des données est parfois mûrement réfléchi, traduisant une volonté plus ou moins prononcée de faciliter leur usage et les répliques possibles de services développés sur plusieurs territoires. Par ailleurs, si le choix du format de données se fait par défaut, en ayant une mauvaise connaissance de la palette des formats ou de leurs usages possibles, les remontées de la communauté des utilisateurs peuvent conduire les collectivités à revoir les formats proposés.

³⁸ <https://developer.jcdecaux.com/#/opendata/vls?page=getstarted>

→ **Concernant les transports** publics urbains, nous pouvons rencontrer des données dans les formats les plus simples (tel que CSV, XLS), ou dans des formats utilisés sous système d'informations géographiques (SHP, KML/ KMZ, TAB, ce qui requiert déjà un minimum de connaissance sur les SIG). Les données d'horaires de transport collectif nécessitent des formats plus complexes. Elles sont diffusées selon le format standard GTFS ou selon la norme Neptune. Les collectivités ayant diffusé initialement ces données horaires au seul format Neptune ont ainsi élargi leur mise à disposition au format GTFS, plus simple.

Comme l'indique Nantes Métropole, le nombre de formats disponibles pour un jeu de données a été multiplié pour essayer de répondre au mieux aux attentes des utilisateurs. Cela peut engendrer des difficultés techniques internes à la collectivité, notamment pour la publication des données multi-formats. Cela a conduit, par exemple, Nantes Métropole à définir des procédures semi-automatiques de publication des jeux de données, à se former à des outils informatiques de traitement des données ou à définir des développements spécifiques. Une autre possibilité visant à minimiser ces difficultés techniques et à satisfaire le plus grand nombre d'utilisateurs est celle, comme le propose la Ville de Paris, de mettre à disposition des données dans des formats bruts ou simples (de type CSV), qui peuvent être lus par tous, mais qui permettent aussi certaines réutilisations.

En somme, les collectivités argumentent non pas tant sur le format de la donnée que sur la facilité que celui-ci propose pour accéder aux données et pour les réutiliser. Si des questions purement techniques se posent (évolution des techniques, des tailles des serveurs ou des systèmes d'exploitation...), les attentes résident davantage dans la mise à disposition de préconisations techniques sur les formats, voire dans une standardisation visant à ne pas proposer un format unique, mais des formats de données adaptés en fonction de l'usage attendu des données. **Le choix du format et sa complexité doivent ainsi être réfléchis en fonction de l'objectif du jeu de données : un format simple pour des données à objectif de transparence de l'action publique et un format qui peut être plus complexe pour des objectifs de réutilisation [informatique] des données.** La complexité du format n'est pas forcément un réel obstacle, à condition de disposer de documentation et d'outils ergonomiques pour manipuler ce format.

→ Apparaissent alors en filigrane certaines difficultés liées aux normes des données, notamment **dans le domaine des transports**. Celles relatives au transport collectif sont relativement bien connues des acteurs. En revanche, les normes relatives à la circulation routière (comme Datex2 pour le recueil, les échanges, et la diffusion de l'information routière) ne sont jamais mises en œuvre sur les plates-formes, ni même citées lors des interviews réalisés.

3.4.2 - Actualisation des données : des processus à mettre en place

Au-delà du choix du format des données, la mise à jour des données mises en partage est une problématique importante. En effet, cette mise à jour est souvent nécessaire pour donner une information fiable. Pour de nombreuses bases de données, l'actualisation est faite en interne par leur producteur, il serait donc dommage de ne pas publier cette mise à jour sur la plate-forme Open Data.

Si cette mise à jour peut être plus simple au sein des directions des services informatiques (souvent en charge de la démarche Open Data), elle est plus complexe pour les autres services et les partenaires, puisque cela peut nécessiter la mise en place de processus spécifiques et un dialogue continu pour aller chercher l'information.

Même si cela est nécessaire pour que les démarches Open Data vivent de manière pérenne, l'actualisation des données reste globalement très chronophage et demande, bien souvent, la mise en place d'un processus ad hoc. Celui-ci doit définir la communication entre les services, en optimisant les moyens humains à mobiliser, et en sécurisant la fiabilité des données.

3.4.3 - Les capacités d'hébergement de l'information : du temps réel à l'historisation et l'archivage

La donnée libérée et sa mise à jour soulèvent, dans le temps, une autre question : celle de la capacité de stockage de l'information et de son accès. Cette problématique apparaît aussi bien pour les données statiques que dynamiques. Il s'agit de deux usages différents des données qui nécessitent une réflexion sur l'hébergement de la donnée.

Selon la communauté urbaine de Bordeaux, les données en temps réel posent un problème particulier : il faut en effet créer une interface de programmation (API pour Application Programming Interface) spécifique pour que les utilisateurs puissent facilement utiliser ces données (sans avoir à télécharger un fichier). Ce type d'API représente une charge financière et technique (temps passé et compétences nécessaires pour créer l'API) qu'il faut prendre en compte. Par ailleurs, ce type de donnée renvoie à une obligation de disposer de l'information en permanence, et pour cela les serveurs d'hébergement des données doivent être très performants. À ce titre, Nantes Métropole exprime un besoin fort d'hébergement des données afin de pallier les difficultés des serveurs locaux des collectivités. Ces interrogations se posent d'autant plus que la nouvelle génération de l'Open Data s'oriente de plus en plus vers les données en temps réel pour répondre aux demandes et besoins de services pour les utilisateurs et les citoyens.

La capacité d'hébergement de l'information se pose aussi par rapport à la possibilité de disposer de l'historique des données mises à jour régulièrement. Faut-il privilégier la donnée « *one shot* » sans trace de son historique ou, au contraire, faut-il pérenniser la trace de la donnée ? Pour la plupart des utilisateurs, des services et des usages connus des données, la donnée la plus récente suffit. C'est ainsi que les données sont écrasées quand une mise à jour est effectuée. Néanmoins, ces données historisées sont d'un intérêt non négligeable, surtout pour les données en temps réel. Notons que toutes ces données (anciennes) effacées demeurent des données publiques auxquelles on doit pouvoir accéder et qu'on doit pouvoir réutiliser.

Mais au-delà de cette possibilité d'historisation des données, la première question réside dans la fiabilité de l'accès aux données et de leur archivage. Comme s'interroge le Grand Lyon, « *quelle est la garantie qu'une donnée publiée sur la plate-forme Open Data reste disponible dans le temps ?* » Ce problème est d'autant plus important que, selon la nature de la donnée, celle-ci peut exiger des mises à jour régulières et fréquentes, notamment pour les données de circulation routière. Dans une perspective d'évolution pour se rapprocher davantage des utilisateurs, des usages et des services, la problématique de l'archivage et de l'historisation se posera donc. Elle renverra, comme dans le cas des données en temps réel, à la nécessité de mettre à disposition des possibilités d'hébergement de l'information (*Cloud services*) au-delà de la solution des serveurs locaux des collectivités.

3.4.4 - Des difficultés à évaluer la qualité de la donnée

La qualité des données est un autre aspect important pour le succès d'une démarche Open Data. Cette notion de qualité recouvre différents concepts tels que l'exactitude, la précision, la fiabilité, la complétude, la fraîcheur de la donnée, l'exhaustivité, etc.

La complétude ou non des jeux de données tient à de nombreux facteurs, dont les problèmes techniques liés aux outils, ou les incompatibilités entre formats numériques. Certaines collectivités effectuent un contrôle qualité de la cohérence entre les données et les métadonnées, mais considèrent que les données sont complètes et réputées qualifiées en amont par leurs partenaires. Par ailleurs, les collectivités peuvent s'appuyer sur la communauté des utilisateurs pour améliorer en retour la qualification et la qualité des données mises à disposition (remontées des utilisateurs via les forums, les réseaux sociaux et le *crowdsourcing*).

Dans d'autres cas, il n'y a pas particulièrement de regard sur la qualité des applications produites, d'une part faute de temps et d'autre part, parce que l'on considère que le mieux est de laisser l'utilisateur final décider. En effet, une application de mauvaise qualité ne sera pas utilisée.

Enfin, pour favoriser l'essor d'une démarche qualité, des collectivités s'appuient sur un référentiel de bonnes pratiques, tel que Opquast publié en juillet 2011³⁹.

La base d'Arrêts de Transport Collectif (ATC) : un exemple d'évaluation de la qualité des données

À la demande du CGDD, le Cerema a mis en place la base ATC⁴⁰, recensant les arrêts de transport collectif des autorités de transport prêtes à diffuser leurs données. À l'été 2014, la base recense plus de trente réseaux. Chacun est intégré dans l'outil libre Chouette de gestion des données d'offre de transport collectif. La qualité des données est alors analysée selon trois angles :

- la conformité au format annoncé (détection d'erreurs de compression, d'encodage, de fichiers erronés...);
- la structuration des données selon le modèle choisi (mauvaise articulation entre arrêts physiques et arrêts commerciaux);
- la précision des données (essentiellement spatiale : arrêts localisés hors voirie, coordonnées nulles...).

39 Le projet Opquast (*Open Quality Standards* – <http://opquast.com>) a pour objectif de faire avancer la qualité web à travers des outils d'évaluation, des ressources, etc. Opquast propose notamment une liste de 72 bonnes pratiques relatives à la qualité pour les producteurs de données Open Data (<http://checklists.opquast.com/fr/opendata>)

40 Détails sur la base ATC et accès à la carte en ligne : www.certu.fr/base-atc-base-nationale-d-arrets-a1217.html

4 - Les clés du montage d'une plate-forme Open Data

4.1 - S'entourer et être soutenu

La plupart des porteurs de projets soulignent **l'importance d'un fort soutien politique** pour initier et mener à bien la démarche Open Data. Si les attentes principales proviennent généralement des services techniques de la collectivité (Paris, Toulouse Métropole...), la démarche s'amplifie très vite avec le soutien des élus. À Bordeaux, la mise en place de la première version de la plate-forme Open Data résulte d'une commande politique passée par le président de la communauté urbaine et le vice-président du conseil général. Ces deux élus, sensibilisés à l'Open Data par la communication faite autour de ce mouvement en 2010, ont souhaité que leurs collectivités se lancent ensemble dans l'Open Data afin de répondre aux enjeux citoyens liés à l'ouverture des données. À Nantes, la démarche Open Data a été initiée par le président de la communauté urbaine, à la suite d'une rencontre avec l'association LiberTIC début 2011.

Dans la plupart des démarches, **des partenariats ont été noués** afin d'enrichir les jeux de données diffusés sur la plate-forme Open Data. Ces partenariats peuvent concerner d'autres collectivités (Région PACA...) ou des acteurs privés comme les exploitants des réseaux de transport en commun (Rennes Métropole, Toulouse Métropole...), les exploitants des services de vélo en libre-service (Toulouse Métropole...), ou encore les exploitants des parkings. Certains de ces partenaires peuvent avoir un rôle moteur dans la démarche de la collectivité, comme l'exploitant du réseau de transports urbains de Rennes. Toutefois, dans une majorité de cas, les collectivités qui portent le projet Open Data doivent engager un travail pédagogique auprès de leurs partenaires pour les convaincre de l'intérêt d'ouvrir leurs données et de les rendre accessibles sur la même plate-forme. C'est particulièrement le cas lorsque les agglomérations engagées dans une démarche Open Data souhaitent impliquer les communes, qui ne sont pas toujours sur la même dynamique et n'ont pas forcément les structures techniques nécessaires pour alimenter le portail (Toulouse, Rennes, Lyon...). Toutes les collectivités interviewées indiquent que l'intégration d'un partenaire à la démarche Open Data se fait de façon pragmatique et non contraignante (les communes doivent ainsi délibérer pour intégrer la plate-forme Open Data de l'agglomération).

L'association peut être plus difficile avec certains acteurs, même si ceux-ci sont déjà engagés dans une démarche Open Data. Ainsi, certains exploitants ont pu refuser de mettre leurs données en partage sur une plate-forme mutualisée pour que leur propre démarche conserve une meilleure visibilité.

Les porteurs de projet Open Data dans les collectivités peuvent souvent compter sur un **appui de la part des associations ou des entreprises locales**. La démarche Open Data de la Ville de Nantes et de Nantes Métropole a cherché à associer le plus possible les acteurs du territoire pour répondre à leurs attentes et bénéficier de leurs conseils. Un partenariat a ainsi été engagé dès le départ avec l'association nantaise LiberTIC qui apporte son expertise et assure un lien permanent avec les différentes communautés (développeurs, journalistes...) souhaitant réutiliser les données publiques libérées. À Montpellier, le montage de la plate-forme de l'agglomération a donné lieu à de nombreux échanges avec les associations et les pépinières d'entreprises, notamment sur leurs attentes. À Toulouse, une lettre d'un collectif citoyens sollicitant un accès aux données a été l'élément déclencheur de la démarche Open Data.

Les porteurs de projet s'appuient également sur les **différents groupes de travail existants sur l'Open Data**. La majorité des collectivités interviewées (Bordeaux, Montpellier, Nantes, Rennes...) ont participé activement au sein du collectif OpenDataFrance⁴¹, dont l'objectif est de mutualiser les bonnes pratiques. Un certain nombre de collectivités qui ne participaient pas au collectif lors du lancement de leur démarche (Lyon, ainsi que des collectivités en dehors de celles interviewées : Lille, Nice, ...) l'ont aujourd'hui rejoint.

Outre OpenDataFrance, les collectivités participent à d'autres réseaux comme le groupe de travail de l'Adullact⁴² sur l'Open Data (Rennes, Toulouse...), ou encore le groupe de travail de l'AFIMB⁴³ sur la réutilisation des données de transport (Rennes, Lyon...). Certains porteurs de projets (Nantes) regrettent de ne pas pouvoir participer davantage à ces groupes de travail, faute de moyens humains.

De façon générale, **les démarches Open Data sont largement alimentées par les retours d'expériences**. L'agglomération de Rennes étant la première collectivité française à s'être lancée dans la démarche Open Data, elle a cherché à s'inspirer des démarches Open Data initiées par d'autres agglomérations étrangères, comme Londres, Boston ou Chicago. À Paris, la démarche Open Data a également été en partie motivée par les démarches similaires lancées par certaines villes nord-américaines et britanniques. La ville et la communauté urbaine de Nantes ont pu bénéficier du soutien de l'agglomération rennaise pour lancer la première version de leur plateforme. Des partenariats forts avaient en effet été tissés dans différents domaines entre ces deux agglomérations, facilitant l'accord de l'agglomération rennaise pour que la plateforme nantaise s'appuie sur le socle technique de leur plateforme. D'autres projets (Toulouse, Région PACA...) ont pu bénéficier du retour d'expérience des premières démarches Open Data lancées en France (Rennes, mais également Paris) pour anticiper les difficultés rencontrées, connaître les points forts et les points faibles de la plateforme, etc.

4.2 - Définir ses attentes pour définir son projet

Le lancement d'une démarche Open Data est motivée par plusieurs raisons. La première généralement citée par les collectivités interviewées (Rennes, Nantes, Paris, Montpellier, département de la Gironde, Région PACA...) est la **volonté de transparence vis-à-vis de la population**, en offrant la possibilité aux citoyens de trouver, sur un portail unique, la plupart des données publiques locales. Certaines collectivités voient également dans la démarche Open Data un moyen de se mettre en conformité avec la directive européenne Inspire, qui impose aux collectivités d'ouvrir les données d'information géographique liées à l'environnement (Lyon).

41 OpenDataFrance (www.opendatafrance.net), collectif créé par des collectivités locales en 2012, a acquis le statut associatif le 9 octobre 2013. Cette association a pour but de regrouper et soutenir les collectivités engagées activement dans une démarche d'ouverture des données publiques et de favoriser toutes les démarches entreprises par ces collectivités dans le but de la promotion de l'Open Data. La quasi-totalité des collectivités interviewées dans le cadre de cette étude sont membres fondateurs de l'association.

42 L'Adullact (www.adullact.org) est une association créée le 24 septembre 2002. Elle a pour objectif de soutenir et coordonner l'action des administrations publiques et des collectivités territoriales dans le but de promouvoir, développer, mutualiser et maintenir un patrimoine commun de logiciels libres utiles aux missions de service public. L'Adullact a lancé un groupe de travail sur l'Open Data en juin 2012.

43 L'Agence française pour l'information multimodale et la billettique (AFIMB) est un service de l'État créé en 2011. L'AFIMB est principalement chargée de promouvoir l'interopérabilité dans les domaines de l'information multimodale et de la billettique, d'encourager le développement de services d'information à l'intention des usagers prenant en compte l'ensemble des modes de transport, et de mettre les systèmes d'information multimodale en réseau. L'AFIMB a réalisé une étude sur l'Open Data dans le domaine des transports [AFIMB 2013].

Une autre raison fréquemment évoquée (Nantes, Toulouse, Paris, département de la Gironde...) est la **création de valeur pour le territoire**, permettant *in fine* d'offrir de nouveaux services à la population. Les collectivités estiment en effet que la transmission, le partage et la réutilisation de ces données dématérialisées favoriseront l'innovation et l'activité : l'objectif est alors de stimuler l'innovation et de développer le tissu économique local. Pour cette raison, certaines collectivités (Nantes, Toulouse...) s'adressent avant tout aux chercheurs et aux développeurs, qui, par le traitement et l'analyse croisée des données, sont en mesure d'y apporter une plus-value. À terme, la mise à disposition d'un plus grand nombre de données économiques, touristiques, ou liées à la vie quotidienne permettrait de renforcer l'attractivité du territoire. Ainsi, les acteurs privés du tourisme pourraient, par exemple, offrir des services plus complets et des informations plus riches à leurs clients.

La démarche Open Data est également perçue comme **un élément positif en termes d'image et d'attractivité du territoire** par plusieurs collectivités (Lyon, Toulouse, département de la Gironde...). L'ouverture des données est un moyen d'animer une dynamique positive rejaillissant sur l'image de la collectivité et de son territoire : la démarche Open Data permet de communiquer sur le fait que la collectivité est sur le terrain de l'innovation, de la transparence, etc. Par ailleurs, les démarches ludiques ou participatives, comme les concours destinés à stimuler la réutilisation des données, ou les initiatives d'enrichissement participatif de la donnée, constituent de nouveaux modes de relations avec les citoyens et les acteurs locaux qui sont perçus favorablement en termes d'image. À l'opposé, plusieurs collectivités admettent que la volonté de ne pas paraître en retard par rapport à d'autres collectivités possédant déjà leur plateforme a été un élément important dans la décision de lancer leur démarche Open Data.

Dans certaines collectivités, **l'existence de projets structurants, partenariaux ou non, nécessite de disposer d'une plate-forme de données**. L'Open Data est donc un outil qui facilite la mise en œuvre de ces projets. À Lyon, le projet Smart Data s'inscrit dans un programme plus large de « métropole intelligente » comprenant une quarantaine de projets (certains s'appuient sur les jeux de données mis en partage). À Rennes, l'agglomération et ses partenaires portent un projet commun de travail en réseau pour proposer de nouveaux services à la population. Cela nécessite de partager de nombreux jeux de données entre les différents partenaires, ce que facilite la plate-forme Open Data.

Une autre finalité régulièrement mise en avant par les collectivités (Paris, Lyon, département de la Gironde...) est **la facilitation des échanges de données au sein de la collectivité** comme avec ses partenaires institutionnels. Les collectivités ont un besoin continu d'échanges d'informations, entre les différents services ou avec leurs différents partenaires. Avant la mise en place du portail Open Data, ces échanges répétés s'avéraient laborieux et chronophages pour les services. La plate-forme Open Data est un outil qui permet de simplifier la logistique de ces échanges et d'économiser du temps. Par ailleurs, certaines collectivités (Paris...) observent qu'après plusieurs mois d'existence la démarche Open Data permet de décloisonner les services et d'améliorer les processus de production. Même si ce gain ne faisait pas nécessairement partie des finalités de la démarche, il fait aujourd'hui partie des résultats appréciés par un certain nombre de collectivités ayant mis leurs données en partage.

Plus rarement, l'Open Data est utilisé comme un outil permettant d'améliorer la qualité de la donnée, via le retour des utilisateurs (*crowdsourcing*). Enfin, la mise en place d'une plate-forme Open Data est également l'occasion pour une collectivité d'inciter ses partenaires à ouvrir leurs données à leur tour.

4.3 - Être pragmatique et progressif dans la mise en œuvre

La durée de montage d'une plate-forme Open Data est généralement relativement brève. Le délai observé entre la prise de décision politique, généralement traduite par une délibération de la collectivité, et la mise en service de la plate-forme est de l'ordre de 6 à 12 mois. Cela montre que les collectivités peuvent se lancer dans une démarche d'ouverture des données sans que le processus ne soit trop fastidieux.

Les collectivités pour lesquelles les temps de montage de la plate-forme sont les plus rapides (Rennes, Toulouse, département de la Gironde, Région Aquitaine...) sont prêtes à mettre en partage certains jeux de données de qualité moyenne, ou à diffuser ces données sur une plate-forme provisoire, et donc à faire face aux critiques des utilisateurs, avant de les faire évoluer en fonction des remontées des utilisateurs.

Ainsi, 8 à 12 mois après le lancement de leur première plate-forme Open Data, plusieurs collectivités (Lyon, Nantes, département de la Gironde...) ont mis en ligne une deuxième version plus évoluée, plus ergonomique, comprenant plus de jeux de données et répondant mieux aux attentes (exprimées ou supposées) des utilisateurs.

Par ailleurs, les temps de gestation du projet Open Data, en amont de la prise de décision politique, peuvent être plus longs. Ainsi, à Toulouse, la première initiative d'ouverture des données est lancée en 2008 sous le nom Toulouse Open. La démarche n'aboutit pas pour des raisons juridiques (choix des licences) et politiques (proximité des élections municipales et création de la communauté urbaine). La démarche est relancée en mars 2011, pour une mise en ligne de la plate-forme en octobre 2011. À Lyon, les premières réflexions sur l'Open Data sont initiées en 2010 pour un lancement du premier catalogue de données en septembre 2012.

4.4 - Disposer des moyens humains et financiers nécessaires

Au sein de la collectivité, le pilotage des démarches Open Data est souvent confié au service en charge de l'informatique, des systèmes d'information ou de l'innovation, qui reçoit l'appui des services juridique et communication. Les autres services sont associés, a minima pour la fourniture et la validation des jeux de données, et pour l'enrichissement de la plate-forme.

L'ensemble de la démarche Open Data est souvent réalisé en interne par la collectivité (Lyon, Toulouse, Paris, première version de la plate-forme du département de la Gironde...). Quelques collectivités font toutefois appel à des prestataires extérieurs, notamment pour la réalisation de la plate-forme internet (Nantes, Rennes, deuxième version de la plate-forme du département de la Gironde...).

Les **coûts d'investissement** liés au développement d'une première version de plate-forme Open Data restent relativement modérés, mais varient assez fortement d'une collectivité à l'autre : ils s'échelonnent de 20 000 euros (Rennes, Toulouse) à 85 000 euros (Montpellier Agglomération). Lorsqu'une nouvelle version de la plate-forme est déployée, généralement après 8 à 12 mois de fonctionnement, on constate que les coûts d'investissements sont comparables à ceux de la première version. Nantes a ainsi consacré 30 000 euros à chacune des deux versions de sa plate-forme, et 45 000 euros supplémentaires à l'évolution de la deuxième version.

À Lyon, l'investissement lié à la première version de la plate-forme est de l'ordre de 60 000 euros, contre 30 000 euros pour la seconde version. Notons que des solutions existent à moindre coût (par exemple pour une première version), en utilisant les plates-formes existantes de niveau supérieur (départemental, régional, voire national avec data.gouv.fr) ou en utilisant des solutions en marque blanche (le site internet de la collectivité n'est qu'une vitrine qui affiche le contenu d'un portail distant, ne nécessitant pas de développement spécifique).

Les **coûts de fonctionnement technique** (hors moyens humains), liés par exemple à l'hébergement des données et à la maintenance de la plate-forme, restent très faibles. Plusieurs collectivités (Toulouse, Région PACA...) n'identifient pas de poste de coût de fonctionnement spécifique lié à l'Open Data, ce coût étant noyé dans la maintenance de l'ensemble du système d'information de la collectivité. D'autres affichent des coûts de fonctionnement de l'ordre de quelques milliers à quelques dizaines de milliers d'euros par an : 5 000 euros par an pour la Ville de Montpellier, et jusqu'à 37 000 euros par an pour la plate-forme mutualisée de Nantes (Ville, communauté urbaine et département) où l'hébergement des données est assuré par un prestataire extérieur.

En revanche, les **coûts de fonctionnement liés à l'animation externe** sont souvent très significatifs. Il faut ainsi compter environ un demi-équivalent temps-plein pour animer une démarche Open Data. À cela il faut ajouter les coûts liés à l'organisation de concours ou d'appels à projets destinés à promouvoir la réutilisation des données ouvertes. Nantes, Rennes ou encore le département de la Gironde ont ainsi consacré plusieurs dizaines de milliers d'euros en prix pour les lauréats (parfois aidés de sponsors).

Les **moyens humains consacrés à la mise en place** de la plate-forme Open Data sont de l'ordre de 1 à 3 équivalents temps-plein, suivant les collectivités. À de rares exceptions près (Nantes...), le lancement de la démarche Open Data ne donne pas lieu à des créations de postes : le travail supplémentaire est généralement absorbé par les équipes en place.

La **maintenance de la plate-forme** mobilise généralement moins d'un équivalent temps-plein. Toutefois, l'animation et la communication nécessaires en interne à la collectivité peuvent nécessiter des moyens humains plus conséquents, qu'il reste difficile d'estimer. Hormis l'équipe projet, généralement resserrée autour de 2 à 3 personnes pour qui l'Open Data constitue une des missions principales, les démarches Open Data s'appuient fréquemment sur des correspondants dans chaque service. Ainsi, à Paris, depuis 2011, un réseau de correspondants issus de chaque direction permet aux services de s'acculturer et de partager les valeurs et la vision de la démarche Open Data. Ce réseau se fixe régulièrement des objectifs chiffrés de publication de jeux de données, relaie les besoins des services et propose les nouveaux jeux de données à partager. À Rennes, l'agglomération organise auprès de ses services des sessions InfoLab sur les thématiques liées à l'Open Data (Quels intérêts ? Quels usages ? Pourquoi ouvrir les données ? Quelles sont les données ouvertes ?). Ces sessions sont un levier pour améliorer le fonctionnement de l'administration et travailler davantage en collaboration autour de la démarche Open Data. L'agglomération de Rennes estime que l'alimentation, la maintenance et l'animation liée à la démarche Open Data mobilisent de l'ordre d'1,5 équivalent temps-plein.

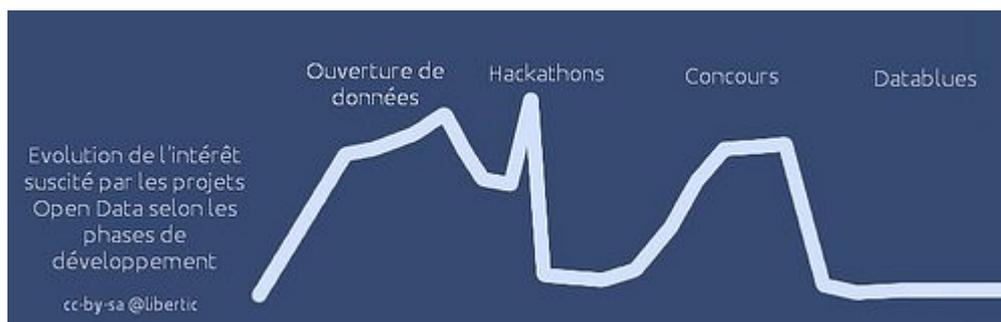


Illustration 6 : Importance de l'animation dans un projet Open Data
(Source LiberTIC, 2013, <https://creativecommons.org/licenses/by-sa/2.0/fr/>)

Plusieurs collectivités (Toulouse, Montpellier, Lyon...) soulignent toutefois que si la démarche Open Data nécessite des moyens humains, elle permet de simplifier la logistique des échanges de données avec les différents partenaires et donc de gagner du temps.

4.5 - Ouvrir rapidement des données sur la mobilité

Les données de mobilité apparaissent souvent comme un thème phare des plates-formes Open Data car, avec les données géographiques, elles font partie des thématiques intéressant fortement les utilisateurs [Serdalab, 2013]. Ainsi, certaines démarches Open Data (Lyon, département de la Gironde, Région PACA) ciblent prioritairement les données liées à la mobilité⁴⁴.

Toutefois, devant la difficulté de mesurer a priori l'utilité de la donnée pour les utilisateurs, un certain nombre de collectivités (Paris, Toulouse, Montpellier) souhaitent commencer par mettre en partage les données les plus faciles à diffuser, sans cibler spécifiquement de thématique particulière. Les données ouvertes sont alors prioritairement les données fiables, produites par la collectivité, et pour lesquelles celle-ci dispose de l'ensemble des droits de propriété intellectuelle.

Concernant ces données relatives à la mobilité, l'un des problèmes fréquemment soulevés relève de la propriété de la donnée. Ainsi, les données des réseaux de transports en commun appartiennent très souvent aux sociétés exploitant ces réseaux, qui craignent de perdre un avantage commercial par rapport à la concurrence en le diffusant. Les collectivités commencent donc, lorsqu'elles passent un contrat avec ces exploitants, à prévoir une clause ouvrant la possibilité de libérer ces données.

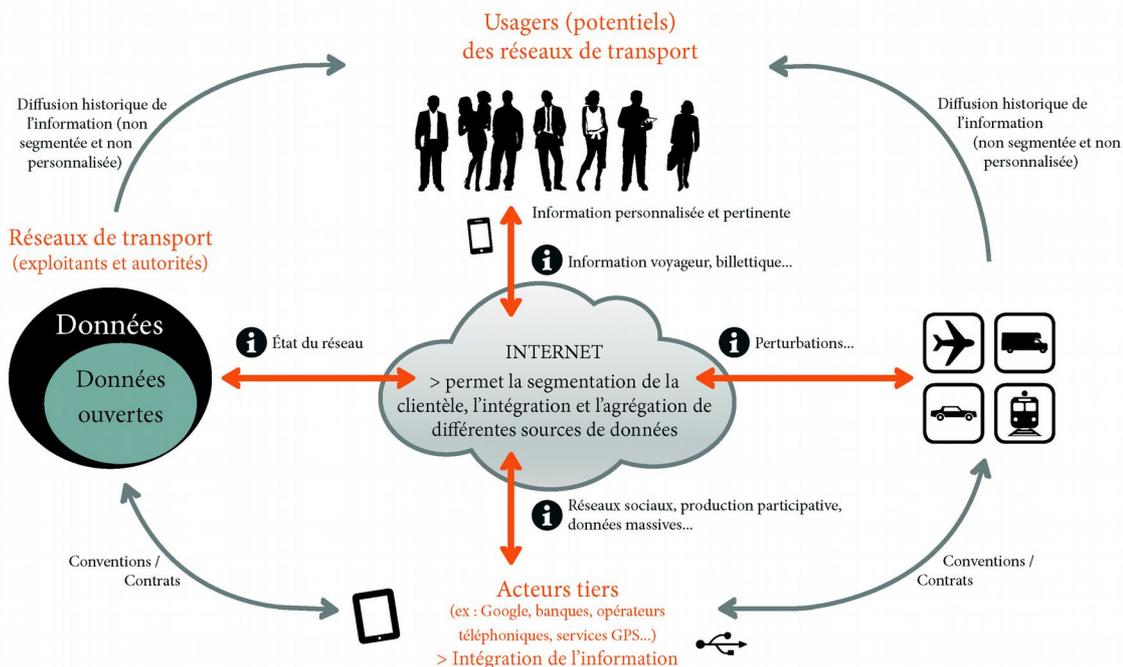
Dans tous les cas, il faut être conscient que la dynamique autour de la démarche Open Data sera plus forte si des données de mobilité y figurent, et particulièrement les horaires des transports collectifs.

Le détail des données liées aux transports et à la mobilité diffusées sur les plates-formes Open Data est indiqué au chapitre suivant.

L'illustration 7 montre l'écosystème d'acteurs autour des données de transport, ainsi que les bénéficiaires potentiels.

⁴⁴ Le détail des avantages à ouvrir des données sur la mobilité figure au chapitre 2.4.

PANORAMA DES DONNÉES OUVERTES DES RÉSEAUX DE TRANSPORT



Qui en profite ?

Les usagers des réseaux de transport

(meilleure qualité de service)

- i** Information personnalisée, partout, tout le temps
- 🕒** Gain de temps grâce à l'optimisation de ses déplacements
- 💰** Économies grâce à une meilleure information
- 😊** Voyage plus agréable en évitant les perturbations

Les réseaux de transport (exploitation plus efficace)

- 👤** Meilleur accès à l'information = plus de passagers
- 💰** Applications développées par des tiers = économies
- 📶** Passagers mieux informés = utilisation optimisée de la capacité du réseau
- 👍** Image positive (transparence et ouverture)

Le secteur économique (nouveaux marchés)

- 💡** plus d'innovation et de concurrence
- 👤💻** Création d'emplois qualifiés
- 🤝** Nouvelles opportunités d'affaires pour les acteurs tiers et les développeurs

© UITP - Avril 2014 - <http://www.uitp.org/benefits-open-data> - Traduction : Cerema

Illustration 7 : écosystème et enjeux autour des données de mobilités

La réflexion nantaise sur les premiers jeux de données

« Au lancement de notre plate-forme, nous avons choisi d'ouvrir principalement des jeux de données pour trois thématiques principales. D'abord la thématique « Culture et Tourisme », car c'est un thème qui intéresse un large public et que la Ville de Nantes et Nantes Métropole voulaient tous deux mettre en avant ; Ensuite la thématique « Environnement », car la ville de Nantes candidatait à l'époque pour être capitale verte européenne, mais aussi car c'est un thème majeur pour tout acteur public depuis le Grenelle de l'Environnement ; et enfin, la thématique « Mobilité », pour plusieurs raisons :

- c'est un des axes forts de la politique de la Ville de Nantes et surtout de Nantes Métropole depuis maintenant plusieurs années ;*
- c'est un des leviers du développement durable (qui, comme le thème « Environnement », est un thème majeur pour les collectivités depuis le Grenelle de l'Environnement) ;*
- LiberTIC (mouvement citoyen sur lequel Ville de Nantes et Nantes Métropole se sont appuyés pour monter leur plate-forme) avait fait remonter des besoins forts des professionnels et des citoyens sur l'ouverture des données liées à la mobilité ;*
- le projet nantais s'est fortement inspiré de la plate-forme rennaise alors déjà en place et très axée sur la mobilité.*

Depuis le lancement de la première version de la plate-forme fin 2011, les nouvelles données mises en partage sont choisies pour répondre aux attentes des citoyens (d'où une nouvelle thématique forte : les données budgétaires) et des développeurs. »

Entretien accordé par Nantes Métropole au Cerema, le 28 mai 2013

5 - Mobilité : le choix des jeux de données

Ce chapitre 5 dresse un panorama des données de mobilité rencontrées, en décrivant précisément les types de données et leur format (par format, on entend le format informatique du fichier, mais aussi la structure interne qui sert de modèle de donnée). Ce panorama permet ensuite d'esquisser des recommandations techniques pour une meilleure réutilisation des données : qualité des données et interopérabilité (le chapitre 6 propose quant à lui des recommandations plus générales sur la mise en œuvre d'une plate-forme Open Data). Les différents jeux de données sont regroupés dans les catégories suivantes, relatives au mode de transport :

- les données de transports collectifs ;
- les données sur les vélos et pistes cyclable ;
- les données routières et de voirie ;
- les données de stationnement ;
- les données relatives aux autres modes de déplacements.

À la fin de ce chapitre 5, une synthèse des recommandations techniques est proposée.

5.1 - Les transports collectifs urbains et non urbains

5.1.1 - Principaux jeux de données sur les transports collectifs

Les principales données relatives aux transports publics qui sont mises en partage par les collectivités ou les AOT sont les données sur la description géographique et sur la qualification de l'offre de service des réseaux. Toutefois, nous trouvons selon les plates-formes des informations complémentaires qui peuvent s'avérer utiles pour les utilisateurs. Ci-dessous, nous en faisons un récapitulatif en distinguant les données relatives aux réseaux urbains, aux réseaux départementaux et aux réseaux régionaux.

→ **Recommandations** : Pour l'information statique relative aux transports collectifs, les données d'offre (arrêts, lignes, horaires, tarifs) n'ont pas vocation à être diffusées au grand public, car des sites internet spécifiques donnent déjà cette information de manière plus ergonomique. Il ne faut pas chercher la transparence mais plutôt la réutilisation pour ces données. Il est donc recommandé d'utiliser prioritairement le format standardisé GTFS, qui peut être utilisé pour développer des applications pour smartphone. Dans la mesure du possible, il est recommandé de doubler ce format GTFS par un format plus complet que sont les normes Neptune et NeTEx⁴⁵, permettant des fonctionnalités plus larges. Pour chaque utilisation de format standardisé, il est recommandé de tester la conformité du jeu de données au format⁴⁶. En revanche, les données liées à l'usage (fréquentation, statistiques, finances) relèvent plus de la volonté de transparence et doivent donc être diffusées dans des formats simples. Enfin, si la collectivité ne diffuse pas les données horaires de son réseau de transport urbain, il est recommandé de mentionner l'existence de ces données (avec les conditions de diffusion) et de diffuser a minima les données géolocalisées relatives aux arrêts⁴⁷.

45 Neptune est aujourd'hui la norme recommandée, mais dès que la norme NeTEx sera officiellement publiée par le CEN, puis par l'AFNOR, il sera recommandé d'utiliser NeTEx, plus complète. Pour plus de précision sur les normes relatives à l'information voyageur, voir : www.normes-donnees-tc.org

46 Pour le standard GTFS, il existe un outil en ligne : *GTFS Feed Validator*. Pour les normes Neptune et NeTEx, il existe l'outil libre Chouette, disponible sur www.chouette.mobi

47 Ces données peuvent ainsi venir intégrer la base nationale d'arrêts de transport collectif réalisée par le Cerema : <http://www.certu.fr/base-atc-base-nationale-d-arrets-a1217.html>

Pour les réseaux de transports urbains, nous trouvons principalement sur les plates-formes Open Data les arrêts, les lignes et leurs caractéristiques, ainsi que les données horaires. Très peu de collectivités gardent en ligne l'historique des mises à jour. Elles sont également peu nombreuses à fournir des données sur la fréquentation des réseaux urbains.

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Arrêts	Localisation des points d'arrêts des lignes des transports collectifs urbains, avec les attributs principaux des arrêts (libellé, sens et nom des lignes passant à l'arrêt) <i>Distinction possible des modes de transports (bus, tramway et métro) pour les réseaux urbains ; bus interurbains et scolaires pour EPCI</i>	SHP (11), CSV (10), GTFS (5), KML / KMZ (4), XLS (3), XML, XML Neptune, Mapinfo, JSON, DXF	Ville de Paris
	Correspondances stations/lignes sur le réseau		Communauté urbaine de Bordeaux, Toulouse Métropole, Montpellier Agglo, Nantes Métropole, Rennes Métropole,
	Accessibilité des arrêts de bus (liste des points d'arrêts des bus)		Région PACA, Île-de-France (RATP), Pigma (Grand Dax, communauté de communes de Montesquieu)
Lignes	Géographie des lignes des transports collectifs urbains, avec les attributs principaux des lignes (libellé, sens et nom de la ligne, informations spécifiques de la ligne) <i>Distinction possible des modes de transports (bus, tramway et métro) pour les réseaux urbains ; bus interurbains et scolaires pour EPCI</i>	SHP (9), GTFS (5), CSV (4), KML / KMZ (3), XML, XML Neptune, DXF	Communauté urbaine de Bordeaux, Toulouse Métropole, Montpellier Agglo, Nantes Métropole, Rennes Métropole
	Couloir de bus (site propre)	SHP, KML / KMZ, CSV	Région Île-de-France (RATP), Pigma (Grand Dax), PACA
Qualification de l'offre de service de transports collectifs	Offre de service des transports collectifs urbains (horaires journaliers, précisions sur les services de semaine, de week-end, etc.)	GTFS (5), XML Neptune, XML, ODS, XLS, HTML	Communauté urbaine de Bordeaux, Toulouse Métropole, Nantes Métropole, Rennes Métropole Région PACA, Île-de-France (RATP),
Fréquentation des transports collectifs	Trafic annuel entrant par station / Information trafic du réseau	XLS, HTML	Île-de-France (RATP), Rennes Métropole

Tableau 3 : Données sur les transports collectifs urbains

Il est à noter que :

- la plate-forme Open Data du Grand Lyon (smartdata.grandlyon.com) ne met pas encore à disposition les données relatives aux transports collectifs urbains gérés par l'AOTU (le Sytral) et l'exploitant Keolis ;
- la plate-forme Open Data de Montpellier Agglomération propose les données géographiques des réseaux de bus urbains et suburbains ;
- la plate-forme de Nantes Métropole propose deux jeux de données à deux dates différentes sur les transports publics (été 2013, hiver 2013). Serait-ce le début de mise en partage des données sans écraser les versions précédentes lors des mises à jour ?
- la plate-forme PACA héberge des données de mobilité d'autres EPCI⁴⁸ (Ville de Digne-les-Bains, syndicat d'agglomération nouvelle Ouest Provence, communauté d'agglomération Arles Crau Camargue, communauté du Pays d'Aix, Marseille Provence Métropole, Pays d'Aubagne et de l'Étoile) ;

48 EPCI : Établissement Public de Coopération Intercommunale

- la plate-forme Aquitaine Pigma⁴⁹ héberge des données de mobilité de différents EPCI (Grand Dax, communauté de communes de Montesquieu).

Pour les réseaux de transports collectifs départementaux, nous trouvons également les arrêts, les lignes et leurs caractéristiques sur les plates-formes Open Data, ainsi que les données horaires. Des données relatives à la fréquentation, la tarification ou les périmètres d'intervention sont rarement mises en ligne (un seul cas – la Loire-Atlantique – sur l'ensemble des acteurs interviewés).

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Arrêts	Localisation des points d'arrêts des lignes régulières des transports départementaux, avec les attributs principaux des arrêts (libellé, sens et nom des lignes passant à l'arrêt) <i>Distinction possible sur les services scolaires, réguliers et autres</i>	SHP (4), KML / KMZ (4), XML (3), CSV (3), GTFS (2), JSON (2), XLS, Mapinfo	Gironde, Loire-Atlantique, PACA (Bouches-du-Rhône), Pigma (Dordogne)
Lignes	Géographie des lignes régulières des transports départementaux, avec les attributs principaux des lignes (libellé, sens et nom de la ligne, informations spécifiques de la ligne) <i>Distinction possible sur les services scolaires, réguliers et autres</i>	SHP (2), XML (2), GTFS (2), KML / KMZ, CSV, JSON	Gironde, Loire-Atlantique, PACA (Bouches-du-Rhône)
Qualification de l'offre de service de transports collectifs	Horaires des lignes de bus réguliers départementales <i>Distinction possible sur les services scolaires, réguliers et autres</i>	GTFS (2), XML, XML Neptune, XLS, CSV	Gironde, Loire-Atlantique
Fréquentation des transports collectifs	Statistiques sur les fréquentations des différentes lignes régulières de transports départementaux	CSV (2), XLS (2), JSON, XML	Loire-Atlantique
	Élèves transportés (répertoriés par lieu de scolarisation et par lieu de résidence)		
Tarification des transports collectifs	Tarifs pour emprunter le réseau de transport du département	XLS	Loire-Atlantique
Périmètre d'intervention	Périmètre des secteurs de transports départementaux	KML / KMZ (6), SHP (6), CSV, XLS	Loire-Atlantique
	Périmètre d'intervention des syndicats intercommunaux de transports scolaires et localisation des sièges des syndicats intercommunaux de transports scolaires		
	Périmètres de transport scolaire : collèges privés et publics, lycées		
Services	Autocaristes départementaux	CSV	PACA (Bouches-du-Rhône)

Tableau 4 : Données sur les transports collectifs départementaux

Il est à noter que :

- le département de Loire-Atlantique détaille de nombreuses données de mobilité sur la plate-forme Open Data ;
- la plate-forme de la Région PACA héberge des données de mobilité du département des Bouches-du-Rhône ;
- la plate-forme aquitaine Pigma héberge des données de mobilité des départements de la Dordogne et de la Gironde.

49 PIGMA est la plate-forme d'Information Géographique Mutualisée en Aquitaine : <http://www.pigma.org>

Pour les réseaux de transports collectifs régionaux, nous trouvons également les arrêts, les lignes et leurs caractéristiques sur les plates-formes Open Data, ainsi que les données horaires. Toutefois, nous trouvons selon les Régions de nombreuses données complémentaires portant sur la fréquentation, la tarification ou la description physique des réseaux.

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Arrêts	Localisation des gares ferroviaires de tous types, exploitées ou non, avec les attributs principaux des arrêts (libellé, sens et nom des lignes passant à l'arrêt)	CSV (2), SHP (2), GTFS	PACA, Pigma (Aquitaine et Poitou-Charentes), Open Data SNCF (Transilien, TER, Intercités)
	Liste des gares de voyageurs telle que définie dans le Document de Référence du Réseau Ferré National (RFN)		
Lignes	Description du statut des lignes exploitées du réseau ferré national en service (RFN)	SHP, GTFS	Pigma (Aquitaine et Poitou-Charentes), Open Data SNCF (Transilien, TER, Intercités)
Qualification de l'offre de service de transports collectifs	Horaires des Transports Express Régionaux (TER) et Lignes Expresses Régionales	XLS (5), CSV (2), JSON (2), XML (2), ODS (2), GTFS (2), TXT, SHP	Pays de la Loire, PACA, Pigma (Aquitaine et Poitou-Charentes), Open Data SNCF (Transilien, TER, Intercités)
	Heures d'ouverture des guichets des gares régionales		
	Évolution du nombre trains km et cars km du réseau régional de transport		
	Vitesse max de ligne,		
Fréquentation des transports collectifs	Évolution du nombre de voyageurs par liaison par mois réseau lignes Transport Express Régionaux (TER)	CSV (5), JSON (4), XLS (4), XML (4), SHP (2)	Pays de la Loire, Pigma (Aquitaine et Poitou-Charentes), Open Data SNCF (Transilien)
	Évolution du nombre de voyageurs par liaison par mois réseau lignes autocars régionaux		
	Statistiques de l'évolution du nombre de voyageurs par liaison par année sur le réseau TER		
	Statistiques de montées-descentes par commune en gare		
	Trafic de marchandises		
Finances	Éléments du compte d'exploitation TER SNCF	CSV, JSON, XLS, XML	Pays de la Loire
Tarification	Statistiques cartes de réduction et tarification sociale (fréquentation)	XLS (4), CSV (4), JSON (3), XML (3)	Pays de la Loire, PACA, Open Data SNCF (Transilien)
	Tarifs Lignes Expresses Régionales		
Système d'information multimodal	Statistiques du site SIM (nombre de visites, nombre de pages vues, nombre d'itinéraires calculés)	CSV, JSON, XLS, XML	Pays de la Loire
Éléments descriptions du réseau ferré	Électrification, point kilométrique, traversée voie piétons, sous-station électrique, pont-route, pont-rail, tunnel, passerelle, passage à niveau, nombre de voies, point d'accès, passage souterrain	SHP	Pigma (Aquitaine et Poitou-Charentes)

Tableau 5 : Données sur les transports collectifs régionaux

Il est à noter que la plate-forme aquitaine Pigma héberge des données de mobilité des Régions Aquitaine et Poitou-Charentes.

5.1.2 - Les données de transports collectifs en temps réel

Les collectivités mettant en partage des données de transports collectifs en temps réel sont très peu nombreuses, comme le montre le tableau n° 6 ci-dessous. Cette rareté s'explique généralement par le refus des opérateurs de fournir les données, pour des raisons de prudence, de désintérêt ou de protection de d'informations considérées comme ayant une forte valeur commerciale.

Thèmes	Données correspondantes	Formats (occurrence)	Fournisseurs
Qualification de l'offre de service de transports collectifs	API Temps Réel (liste des lignes, des arrêts, des horaires en temps réel)	API (3)	Toulouse Métropole, Nantes Métropole, Rennes Métropole
	Trafic prévisionnel	API	Nantes Métropole
État des réseaux	Commande permettant de récupérer les perturbations, l'état (disponible/indisponible) des équipements, l'état (ouvert/fermé) des arrêts (métro, tramway), les états des parcs relais et points de ventes	API	Rennes Métropole

Tableau 6 : Données de transports collectifs en temps réel

Les données sont diffusées via des API (interfaces de programmation), évitant des mises à jour fréquentes de fichiers volumineux⁵⁰. Les trois exemples connus ci-dessus utilisent chacun un formalisme spécifique, sans s'appuyer sur une norme existante (la norme SIRI par exemple). Les réutilisations sont donc plus limitées car elles ne peuvent fonctionner que sur une seule agglomération.

Malgré cela, les collectivités sont nombreuses à affirmer lors des entretiens vouloir développer leur plate-forme en intégrant cette dimension des données en temps réel. Cela correspond également aux besoins grandissants des utilisateurs. 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 devraient être en mesure de diffuser ces données à moyen terme.

→ **Recommandations** : Pour l'information en temps réel relative aux transports collectifs, la première recommandation est d'ouvrir les données, en raison de leur forte valeur ajoutée et des services innovants qu'elles permettent. Mais pour des services vraiment innovants, une standardisation des formats est nécessaire. Cette standardisation peut se faire en utilisant la norme SIRI-Lite, ou un autre format unique. La première solution nécessite un travail de communication et d'accompagnement auprès des collectivités sur les possibilités offertes par SIRI-Lite et sur la manière concrète de l'utiliser. La deuxième solution nécessite une mise en synergie des collectivités impliquées dans la diffusion de ce type d'information en temps réel, afin de choisir un format standard, sur la base des potentialités et de la facilité pour les collectivités d'exporter leurs données dans ce format. Dans ce cas, le choix d'un API est recommandé, plutôt qu'un fichier complet qui serait mis à jour très régulièrement (toutes les minutes par exemple).

⁵⁰ La plate-forme de la communauté urbaine de Bordeaux propose toutefois une mise à jour de fichiers toutes les cinq minutes pour certaines données, mais cette méthode ne serait pas pertinente pour les transports collectifs car elle nécessiterait une fréquence de mise à jour de quelques secondes.

5.2 - Le vélo et les pistes cyclables

Les principales données relatives aux vélos qui sont mises en partage par les collectivités portent sur les pistes cyclables (objets géographiques) et le mobilier (stationnement des vélos). Toutes les agglomérations ou communes ayant un service de vélos en libre-service ont aussi mis en partage des données, statiques la plupart du temps. Elles proposent également des données en temps réel si le prestataire de service le propose lui-même (la collectivité héberge ou référence les données).

D'autres données moins courantes apparaissent, comme la tarification des vélos en libre-service ou la liste des loueurs de vélos.

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Pistes cyclables	Export de la couche vélo d'Openstreetmap	OSM	Gironde
	Pistes cyclables communales, départementales ou d'agglomération	KML / KMZ (7), SHP (5), CSV, JSON, ATOM, WMS, Mapinfo, XML	Ville de Bordeaux, Ville de Montpellier
	Objet linéaire décrivant une infrastructure dédiée à la circulation des vélos. Il est caractérisé par des attributs (emplacement, description, nom, type d'aménagement...)		Grand Lyon, Toulouse Métropole, Montpellier Agglomération, Nantes Métropole Gironde, Loire-Atlantique
	Plan modes doux (objet linéaire représentant un itinéraire du schéma directeur Plan modes doux)	SHP, KML / KMZ	Grand Lyon
Itinéraires touristiques (longueur de sentier...)	XLS, CSV	Loire-Atlantique	
Stationnement des vélos	Emplacement des arceaux publics de stationnement des vélos (y compris en gare SNCF)	CSV (5), JSON (3), XLS (3), KML / KMZ (2), SHP (2), XML, ATOM, WMS, WFS, DWG, ODS	Ville de Bordeaux, Marseille, Paris Pays de la Loire, PACA Communauté urbaine de Bordeaux
	Capacité des abris à vélos dans les gares et nombre d'emplacements dans les abris collectifs sécurisés	CSV, JSON, XLS	SNCF Open Data (Île-de-France)
	Localisation des emplacements de stationnement des vélos en libre-service (station, terminal de paiement...)	SHP (7), CSV (7), KML / KMZ (5), JSON (3), XLS (2), Mapinfo, XML, WMS, WFS, DWG, DXF	Grand Lyon, Toulouse Métropole, Nantes Métropole, communauté urbaine de Bordeaux, Marseille Provence Métropole, Paris et communes d'Île-de-France ⁵¹ , Rennes Métropole, JCDecaux ⁵²
Tarification	Liste des tarifs applicables dans les stations de vélo en libre-service	CSV (2), XLS, SHP, KML / KMZ, WMS, WFS, DWG	Nantes Métropole, communauté urbaine de Bordeaux
Offre de services	Liste des loueurs de vélo, VTT et vélo électriques	CSV	PACA (Bouches-du-Rhône)

Tableau 7 : Données sur les vélos et pistes cyclables (données statiques)

51 La liste des communes d'Île-de-France où les données des vélos en libre-service (produites par JCDecaux, disponibles sur <http://data.iledefrance.fr>) sont ouvertes est la suivante : Arcueil, Aubervilliers, Bagnolet, Boulogne-Billancourt, Charenton, Clichy, Fontenay-sous-Bois, Gentilly, Issy-les-Moulineaux, Ivry, Joinville, Le Kremlin-Bicêtre, Le Pré-Saint-Gervais, Les Lilas, Levallois-Perret, Malakoff, Montreuil, Montrouge, Neuilly, Nogent, Pantin, Paris, Puteaux, Saint-Cloud, Saint-Denis, Saint-Mandé, Saint-Maurice, Saint-Ouen, Suresnes, Vanves, Vincennes, Cergy, Courdimanche, Éragry-sur-Oise, Neuville-sur-Oise, Pontoise, Saint-Ouen-L'Aumône, Vauréal et Créteil.

52 Sur le site JCDecaux, on trouve les données des vélos en libre-service de Amiens, Besançon, Lyon, Marseille, Mulhouse, Nancy, Nantes, Rouen, Toulouse (et les communes d'Île-de-France citées dans la précédente note de bas de page).

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Stationnement des vélos	Données sur l'état de disponibilité en temps réel des stations de vélo en libre-service (nombre de bornes libres, nombre de vélos disponibles...)	API (4), WMS, WFS, HTML	Communauté urbaine de Bordeaux, Grand Lyon, Rennes Métropole, Toulouse Métropole, JCDecaux
Offre de services	Liste des loueurs de vélo, VTT et vélo électriques	API	PACA (Bouches-du-Rhône)

Tableau 8 : Données sur les vélos et pistes cyclables (données temps réel)

→ **Recommandations** : Concernant les véloroutes et voies vertes, il existe un standard proposé par la COVADIS⁵³. Ce standard pourrait être plus utilisé et, dans cette lignée, un format standard pourrait être défini pour l'ensemble des voies cyclables, notamment en milieu urbain. Pour la disponibilité des stations de vélos en libre-service, il existe un standard propriétaire, proposé par l'exploitant JCDecaux, mais un standard ouvert permettrait de disposer d'un format unique pour l'ensemble des agglomérations, au-delà du choix de l'exploitant.

5.3 - Le stationnement (hors vélo)

De nombreuses données sont mises en partage sur le stationnement automobile et sur le stationnement d'autres véhicules. Les principales informations sont relatives au stationnement public ou sur voirie, au stationnement pour les personnes à mobilité réduite et aux parcs relais. Des données sur les nouveaux services à la mobilité (covoiturage et auto-partage) commencent à voir le jour. Enfin, les données en temps réel ne sont pas la règle, même si plusieurs collectivités interviewées commencent à les mettre à disposition sur leur plate-forme Open Data.

Les données sur les emplacements des camping-cars, des taxis, des autocars ou sur les aires de livraison sont nettement moins abondantes que pour le stationnement automobile.

53 COVADIS : Commission de validation des données pour l'information spatialisée (pour les standards existants, voir : <http://www.certu.fr/les-geostandards-de-la-covadis-r158.html>)

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Stationnement automobile public	Objet ponctuel représentant un parc de stationnement. Il est caractérisé par des informations de gestion (nom, type de parking, nombres de places, type d'usage et de tarification, voies de desserte, gabarits...)	CVS (10), KML / KMZ (9), JSON (7), SHP (6), XLS (4), ATOM (2), Mapinfo (2), XML (2), WMS, WFS, DXF / DWG	Ville de Bordeaux, Montpellier, Nantes, Paris Grand Lyon, Toulouse Métropole, Marseille Provence Métropole, Rennes Métropole PACA (Bouche-du-Rhône)
	Parkings publics hors voirie		
	Liste et position des voies en stationnement payant		
	Localisation des horodateurs		
	Zones bleues		
Stationnement automobile privé et concédé	Objet ponctuel représentant un parc de stationnement. Il est caractérisé par des informations de gestion (nom, type de parking, nombres de places, type d'usage et de tarification, voies de desserte, gabarits...)	WMS, WFS, KML / KMZ, SHP, JSON, CSV	Paris Grand Lyon
Parcs relais	Position des parcs relais avec les attributs (capacité, lignes de TC en correspondance, type d'usagers, nombre de places réservées aux deux-roues, aux PMR, aux véhicules électriques...)	KML / KMZ (3), SHP (3), CSV (3), WMS (2), WFS (2), JSON (2), XML, XLS, DWG	Communauté urbaine de Bordeaux, Grand Lyon, Nantes Métropole, Marseille Provence Métropole
Stationnement pour les personnes à mobilité réduite (PMR)	Liste et emplacement des places de stationnement réservées aux personnes handicapées	KML / KMZ (6), CVS (5), SHP (5), XLS (3), JSON (2), ODS (2), DXF / DWG (2), ATOM, WMS, WFS, Mapinfo	Ville de Bordeaux, Montpellier, Marseille Communauté urbaine de Bordeaux, Toulouse Métropole, Nice Côte d'Azur, Rennes Métropole
Stationnement pour auto-partage	Objet ponctuel représentant un emplacement réservé au stationnement des véhicules d'autopartage. Il est caractérisé par des informations de gestion (adresse, type d'autopartage, nombre d'emplacements)	KML / KMZ (4), JSON (4), SHP (4), CSV (3), XLS (2), WMS, WFS, Mapinfo, XML, DXF / DWG	Grand Lyon, Toulouse Métropole, Nantes Métropole, Rennes Métropole Île-de-France (Paris)
Aire de covoiturage	Objet ponctuel représentant un parc de stationnement. Il est caractérisé par des informations de gestion (nom, type de parking, nombres de places, type d'usage et de tarification, voies de desserte, gabarits...)	JSON (3), KML / KMZ (3), SHP (2), CSV (2), XML (2), XLS (2), WMS, WFS	Grand Lyon, Nantes Métropole Loire-Atlantique
Stations avec bornes de recharge électrique	Localisation des stations de recharges pour les voitures électriques sur le territoire. Ce sont les bornes sur le domaine public (stations mixtes autopartages ou stations multistandards tout public)	CSV (2), KML / KMZ (2), JSON, SHP, Mapinfo, DXF / DWG	Toulouse Métropole, Rennes Métropole
Réglementation	Liste des éléments de circulation et de stationnement réglementés par arrêté municipal	CVS, JSON, ATOM	Ville de Bordeaux
Tarification	Ensemble des tarifs de stationnement	CSV (2), PDF, JSON, XLS, XML	Ville de Montpellier, Nantes, Paris
Temps réel	Disponibilité des parkings en temps réel (localisation et nombre de places)	WMS (2), WFS (2), API (2), KML / KMZ, SHP, JSON, HTML	Communauté urbaine de Bordeaux, Grand Lyon, Nantes Métropole, Rennes Métropole
	Objet ponctuel décrivant un parc de stationnement ouvert au public, lié à des informations d'occupation temps réel et permettant l'affichage de cartographies dynamiques. Il est caractérisé par des informations de gestion (nom, gestionnaire, nombre de places...)		

Tableau 9 : Données sur le stationnement automobile (statiques et temps réel)

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Stationnement pour camping-cars	Aire de camping-cars (stationnement et services, nom)	CSV (3), SHP	Aquitaine, PACA (Bouche-du-Rhône) Gironde
Aires de livraison	Localisation des aires de livraison	CSV (4), KML / KMZ (4), SHP (3), JSON (2), XLS (2), XML(1), Mapinfo, DXF / DWG	Ville de Montpellier, Nantes, Paris Toulouse Métropole, Rennes Métropole
Stationnement des taxis	Localisation des emplacements taxis	KML / KMZ (3), CSV (3), SHP (2), JSON, Mapinfo, XLS (2), DXF / DWG	Ville de Montpellier Toulouse Métropole, Rennes Métropole PACA (Marseille)
Stationnement des autocars et bus (dont bus tourisme)	Liste des parkings autocars départementaux	CSV, DXF / DWG, KML / KMZ, SHP	Rennes Métropole PACA (Bouche-du-Rhône)

Tableau 10 : Autres données de stationnement (statiques)

C'est à Nantes qu'on peut trouver le plus de données ouvertes liées au stationnement. On compte neuf jeux de données sur ce thème, ce qui a permis de mettre en place de nouveaux services aux habitants par le développement de plusieurs applications pour smartphone. Parmi les 27 applications référencées sur la plate-forme Open Data, environ un tiers porte sur le stationnement (un deuxième tiers concerne les différents modes de transport, parfois couplés, et un troisième tiers ne concerne pas le transport)⁵⁴. Le succès de ces données est grandement lié à la mutualisation des différents acteurs qui permet d'avoir une vision globale de cette thématique (parkings de la ville, stationnement sur voirie, parcs relais, aires de covoiturage du département...). Les données relatives au stationnement ont donc un vrai potentiel de réutilisation.

→ **Recommandations** : La première recommandation concernant le stationnement est de viser une diffusion complète de l'offre de stationnement automobile. Ceci implique d'héberger ou de référencer les données d'acteurs autres que le gestionnaire de la plate-forme Open Data. De plus, il n'existe pas aujourd'hui de norme ou de format standard pour décrire cette offre de stationnement. Une telle standardisation ne nécessiterait pas forcément la mise en place d'une norme, mais pourrait se faire par les acteurs légitimes sur le sujet (par exemple : la FNMS⁵⁵, la Féduco⁵⁶, le Gart⁵⁷...).

54 Voir illustration n° 4, page 18.

55 FNMS : Fédération Nationale des Métiers du Stationnement.

56 Féduco : Fédération nationale Du Covoiturage.

57 GART : Groupement des Autorités Responsables de Transport.

5.4 - Le réseau routier

Les compétences en matière de voirie sont réparties entre l'État et les collectivités locales.

À l'échelle nationale, le ministère chargé des transports met des données d'information routière à la disposition des médias et des opérateurs qui élaborent des services à valeur ajoutée. Il s'agit de données sur le trafic et les événements routiers (accidents, bouchons...).

Ces données concernent surtout le réseau routier national, uniquement pour sa partie non concédée, essentiellement au niveau des voies structurantes d'agglomérations (Bordeaux, Caen, Clermont-Ferrand, Grenoble, Lille, Limoges, Lyon, Marseille, Nantes, Paris, Rennes, Saint-Étienne, Strasbourg, Toulouse et la Savoie) et pour quelques cas sur des axes autoroutiers non concédés (A 20 Châteauroux-Limoges, A 75 Clermont-Ferrand-Béziers...). L'État met également à disposition les données de certains départements. Le site www.diffusion-numerique.info-routiere.gouv.fr décrit les données mises à disposition et indique comment les interpréter [Cete de Lyon, 2013]. Les données collectées sur le réseau routier national concédé appartiennent quant à elles aux sociétés concessionnaires et ne sont pas communiquées directement au grand public.

Concernant les collectivités, les données les plus fréquemment ouvertes concernent la description du réseau routier (linéaire de voie, nomenclature des voies, etc.). Les données concernant l'accidentologie (nombre, gravité et localisation des accidents) et les trafics moyens annuels sont diffusées par plusieurs collectivités. Les données en temps réel sont nettement plus rares.

L'Open Data en collectivité à la lumière des données de mobilité

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Description du réseau	Tronçon de voirie	SHP (5), KML / KMZ (4), CSV (3), WMS (2), WFS (2), Mapinfo, DWG, JSON	Grand Lyon, Toulouse Métropole, PACA (Nice Métropole), Nantes Métropole, Loire-Atlantique, Gironde (Brocas), communauté urbaine de Bordeaux
	Nomenclature des voies	CSV (3), SHP (2), KML / KMZ (2), WFS (2), WPS, WMS, JSON	Grand Lyon, Toulouse Métropole, Paris, communauté urbaine de Bordeaux
	Emplacement des panneaux de signalisation fixes	SHP	Ville de Paris
	Emplacement des panneaux à message variable	SHP, CSV, JSON, XLS, XML, KML / KMZ	Nantes Métropole
	Emplacement du mobilier urbain	SHP, CSV, JSON, KML, ATOM	Ville de Paris, Bordeaux
Comptages	Trafic (temps réel)	SHP, KML, WMS, WFS, JSON, XML	Grand Lyon
	Trafic moyen annuel	CSV (3), XLS (2), SHP, KML / KMZ, JSON, XML, ODS	Loire-Atlantique, Ville de Montpellier, communauté urbaine de Bordeaux
Vitesse	Temps de parcours (temps réel)	CSV, JSON, XLS, XML, API	Loire-Atlantique (pont de Saint-Nazaire uniquement) Nantes Métropole
Taux d'occupation	Fluidité des axes routiers	API, SHP, KML / KMZ, WMS, CSW, DWG	Nantes Métropole communauté urbaine de Bordeaux
Zones 30, zones apaisées	Linéaire de zone 30, zone apaisée	KML / KMZ (2), SHP, Mapinfo	Grand Lyon, Toulouse Métropole
Chantier perturbant	Emprise des chantiers perturbant la circulation	KML / KMZ (3), JSON (3), SHP, Mapinfo, CSV, ATOM	Grand Lyon, Toulouse Métropole, Ville de Bordeaux
Accidentologie	Localisation des accidents, des blessés ou des tués sur le réseau routier	CSV (2), SHP, KML / KMZ, XLS, JSON, XML	Loire-Atlantique, Gironde (Brocas)
Incidents	Événements routiers en cours : accidents, déviations... (temps réel)	CSV, JSON (2), KML, SHP, WMS, WFS, XLS, XML (2)	Loire-Atlantique, Grand Lyon
	Ouverture ou fermeture des voies (temps réel)	WMS	Communauté urbaine de Bordeaux (pont Chaban-Delmas uniquement)

Tableau 11 : Données routières diffusées par les collectivités interviewées (statiques et temps réel)

En revanche, les données routières ouvertes par les collectivités ne sont pas formatées selon les normes internationales en vigueur dans le domaine de l'information routière. Les données des différentes collectivités sont hétérogènes, non en termes de standards de l'informatique, qui sont largement partagés, mais en termes de formats des jeux de données et des champs qui les composent. Ce manque d'harmonisation entre les différentes plates-formes Open Data représente un fort verrou technique pour les développeurs qui souhaiteraient créer des applications à l'échelle nationale [Cete de Lyon, 2013].

Par ailleurs, le domaine de l'information routière est déjà relativement bien structuré et concurrentiel. Il existe déjà des fournisseurs d'information trafic et services associés qui utilisent un nombre important de sources de données de trafic (publiques ou privées) qui n'ont pas attendu l'Open Data. Les possibilités de voir émerger de nouveaux services grâce à l'Open Data sont donc moins élevées que dans les autres domaines.

Enfin, il existe un standard pour décrire le réseau des routes à grande circulation, proposé par la COVADIS⁵⁸.

→ **Recommandations** : Le réseau routier mériterait d'être décrit selon un format harmonisé et interopérable. En effet, les informations des collectivités sont généralement plus complètes et fiables que les données couvrant l'ensemble du réseau français ou international. Néanmoins, ces données sont peu utilisées en raison du manque de standardisation de ces formats et de la difficulté à interfacer ces données avec le réseau routier d'autres exploitants. Outre l'harmonisation des formats, c'est aussi d'un processus de juxtaposition et de raccordement des jeux de données dont les réutilisateurs auraient besoin. Ce processus étant complexe, un acteur local (agglomération, plate-forme régionale d'information géographique...) ou national (IGN, OpenStreetMap...) pourrait ainsi être désigné pour réaliser lui-même le raccordement et fournir en sortie un réseau multi-gestionnaires sur un territoire donné.

Pour les informations en temps réel sur le trafic, la norme Datex 2 est conseillée, répondant pleinement aux besoins et permettant une réutilisation facilitée grâce à l'adoption par tous les gestionnaires d'un format unique.

58 COVADIS : COMmission de VALidation des Données pour l'Information Spatialisée (pour les standards existants, voir : <http://www.certu.fr/les-geostandards-de-la-covadis-r158.html>)

5.5 - Les données relatives aux autres modes de déplacement

Quelques plates-formes Open Data diffusent des jeux de données supplémentaires relatifs aux autres modes de déplacements : sentiers pédestres ou informations sur les bacs et les transports fluviaux. Les données diffusées sont statiques.

Thèmes	Jeux de données correspondant	Formats (occurrence)	Fournisseurs
Sentiers pédestres	Linéaire de sentiers pédestres	KML / KMZ (3), SHP, JSON, WMS, PDF	Grand Lyon, Loire-Atlantique
Réseau fluvial	Linéaire de voies navigables	KML / KMZ, SHP	Loire-Atlantique
	Localisation des écluses	KML / KMZ, SHP, XLS, CSV, JSON, XML	Loire-Atlantique
	Horaires des écluses	CSV, JSON, XLS, XML	Loire-Atlantique
	Fréquentation des bacs	CSV, JSON, XLS, XML	Loire-Atlantique
	Localisation des embarcadères des bacs	CSV, JSON, XLS, XML	Loire-Atlantique

Tableau 12 : Données relatives aux autres modes de déplacements

→ **Recommandations** : Pour les sentiers pédestres, s'il s'agit d'un sentier du littoral, le standard de la COVADIS sur ce sujet est recommandé.

5.6 - Recommandations techniques sur les données de mobilité

Les données de mobilité sont déjà bien présentes sur les plates-formes Open Data existantes, mais pour que celles-ci soient sources d'innovation, elles doivent répondre à des exigences accrues d'ouverture et de standardisation, comme l'indiquent les recommandations relatives à chacun des modes⁵⁹. Il convient ainsi d'**ouvrir a minima des jeux de données « classiques »** (arrêts de transport collectif, linéaire routier, emplacements des parkings...) et si possible des données à plus forte valeur ajoutée (horaires de transport collectif, théoriques voire en temps réel; trafic routier...). Il convient également de **standardiser les formats des données** : il faut donc créer des formats standardisés⁶⁰ (parkings, voies cyclables...), rendre ouverts certains standards propriétaires (vélos en libre-service) ou encore utiliser plus largement les normes et standards existants (SIRI ou Datex 2).

Ouverture et standardisation sont des pré-requis, mais doivent s'accompagner d'un souci de la **qualité des données**, qui nécessite la prise en compte de plusieurs aspects :

- la précision géographique ;
- la précision temporelle, qui passe par des mises à jour fréquentes des données ;
- la complétude, qui traduit le fait que l'ensemble des données nécessaires sont disponibles ;
- la cohérence entre les différents jeux de données ;
- la profondeur des données, qui traduit la richesse des informations données sur chaque objet ;
- le renseignement des méta-données, qui informe sur le contenu des données et leur qualité.

⁵⁹ Voir les parties 5.1 à 5.5.

⁶⁰ Les formats standardisés désignent les normes et les standards, comme explicité dans la fiche *Normes et standards de l'information voyageur* réalisée par le Cerema : <http://www.certu-catalogue.fr/catalogsearch/result/?q=normes>

Tous ces aspects sont importants, mais il est recommandé d'être particulièrement attentif à la complétude et à la cohérence. En effet, bon nombre de collectivités ne diffusent que leurs propres données, ce qui est très compréhensible. Néanmoins, pour le réutilisateur, ces données sont peu (ou pas) utiles lorsque les autres gestionnaires ne diffusent pas ce même type de données, ou dans des modalités ou formats différents. Ainsi, la plate-forme nantaise propose l'ensemble des parkings de la ville : ceux gérés par des exploitants privés, ceux gérés par la Ville, les parcs relais gérés par l'agglomération ou les aires de covoiturage gérées par le département. C'est bien grâce à la complétude de l'information relative au stationnement que de nombreuses réutilisations en sont faites.

Le rapprochement de plusieurs jeux de données dont la source est différente amène ensuite la question de la cohérence. Il s'agit non seulement d'une cohérence globale (deux informations contradictoires) mais surtout d'une cohérence dans le référentiel utilisé. Ainsi, les exploitants routiers diffusent généralement les données de leur linéaire routier, selon leur propre référentiel. Certains sont basés sur le réseau IGN, d'autres sur TéléAtlas ou encore OpenStreetMap...

Mais la compilation de deux jeux de données basés sur des référentiels différents se heurte à deux difficultés :

- le décalage géographique entre les objets représentant une même réalité (dû aux données sources ou à la méthode de géolocalisation) ;
- la différence de structure du fichier (attributs, signification des valeurs de ces attributs...). Ce point peut en partie être levé par l'utilisation d'un modèle de donnée commun, qui peut être issu de la norme européenne GDF⁶¹.

Ce même problème empêche souvent le rapprochement entre un réseau de transport collectif et le réseau routier.

Pour aller plus loin dans les recommandations sur les données de mobilité, on pourra se référer à l'étude réalisée en 2013 pour l'AFIMB, avec une approche plus technique et plus pointue [AFIMB, 2013].

61 GDF : *Geographic Data File* (plus de précision sur le site : http://en.wikipedia.org/wiki/Geographic_Data_Files)

6 - Recommandations pour la mise en œuvre d'une plate-forme Open Data

Le chapitre précédent a esquissé des recommandations techniques relatives aux différentes catégories de données de mobilité. Dans ce chapitre, il s'agit de proposer des recommandations plus globales sur la mise en œuvre d'une plate-forme Open Data.

6.1 - Commencer l'aventure : être pragmatique

Le lancement d'une démarche Open Data pose un grand nombre de questions d'ordre technique, juridique, organisationnel, etc. Par exemple, les collectivités qui souhaitent lancer une démarche Open Data s'interrogent souvent sur la qualité des données diffusées : faut-il diffuser des données dont la qualité est jugée insuffisante ? Le choix des licences fait généralement l'objet de débats au sein des collectivités : faut-il choisir la licence ouverte d'Etalab, la licence ODbL ou encore créer une licence « sur mesure » ?

Ces questionnements sont parfaitement légitimes, mais peuvent ralentir fortement la démarche Open Data. Le retour d'expérience des collectivités interviewées montre qu'il est préférable de se montrer pragmatique et d'avancer de façon progressive. Il est préférable de ne pas essayer de diffuser dès le lancement de la plate-forme l'ensemble des jeux de données avec toutes les fonctionnalités, mais plutôt d'avancer progressivement vers une offre de plus en plus complète. Au niveau qualitatif également, mieux vaut commencer avec des données existantes (ou ne nécessitant qu'un faible travail d'amélioration de la qualité), sans chercher à avoir des données parfaites dès le départ. Pour les licences, il est enfin préférable de choisir une des deux licences majoritaires : licence ouverte ou licence ODbL, garantissant ainsi le cadrage juridique sans avoir à lancer un travail conséquent.

Le choix des jeux de données initiaux : l'exemple de la ville de Paris

« Nous avons commencé par les données utiles et faciles à mettre en partage, sur la base des critères suivants :

- *accessibilité de la donnée (facilité à exporter) ;*
- *absence de contrainte juridique (propriété intellectuelle) ou politique ;*
- *intérêt de la donnée pour le citoyen.*

La volonté de la Ville de Paris était de mettre un maximum de données en partage (respectant ces trois critères) afin d'avoir un retour important de la part des utilisateurs. La Ville souhaitait en effet faire évoluer sa plate-forme à partir de ces remontées. Il était donc important, pour susciter les réactions et créer une émulation, de mettre un certain nombre de données en ligne, et parmi elles, des données intéressant le grand public (comme le jeu de données sur les prénoms les plus courants à Paris, qui est le jeu de données le plus consulté). C'est grâce à ce feedback que nous améliorons la plate-forme et que la seconde version, mise en ligne le 21 mai 2013, a été pensée.

Nos agents en interne sont parfois réticents à mettre ces données de qualité moyenne en ligne, car ils ont peur des mauvaises remontées et aussi peur de devoir répondre à de nombreuses questions. Mais si le jeu de données n'est pas parfait, c'est que nous n'avons pas besoin, en interne, qu'il soit de meilleure qualité pour faire notre travail. Il n'y a pas à rougir de ça. Les remontées des utilisateurs nous permettent désormais d'améliorer ces jeux de données. »

6.2 - Bénéficiaire d'un portage politique volontaire et fort

L'ensemble des collectivités interviewées s'accordent à dire qu'un portage politique fort est indispensable pour asseoir la légitimité du projet et permettre un lancement rapide de la plateforme Open Data. Le portage politique est nécessaire pour faire adhérer l'ensemble des services de la collectivité à la démarche Open Data, qui peut parfois être perçue comme très technique et remettant en cause les pratiques de travail au sein de la collectivité. À ces fins, il peut être utile de présenter cette démarche Open Data auprès des élus et directeurs les plus volontaristes de la collectivité ou de ses partenaires. Cette présentation doit montrer les résultats tangibles déjà obtenus, écouter les doutes et les inquiétudes, afin de rassurer en apportant des réponses et des propositions.

6.3 - Convaincre et communiquer

Un exercice de pédagogie est ensuite nécessaire, d'abord en interne à la collectivité, auprès des agents, pour les convaincre de l'intérêt de l'Open Data et les faire adhérer à la démarche. L'accompagnement des agents, par exemple par la mise en place de correspondants Open Data dans les différents services, facilite l'appropriation de la démarche et permet de faire de l'Open Data une logique « naturelle » de valorisation des travaux des agents. Néanmoins, cette adhésion se fait dans la durée. Au démarrage du projet, il est préférable de commencer par faire adhérer les services et partenaires qui sont les plus moteurs et les plus sensibles à ce genre de projets. L'adhésion pourra ensuite plus facilement se répandre dans les autres services, surtout lorsque la plate-forme sera mise en place et que chacun pourra mesurer le résultat concret du projet.

6.4 - Profiter des réseaux des collectivités et de la communauté de l'Open Data

Le retour d'expérience des démarches Open Data des collectivités commence à devenir fourni : il constitue une matière intéressante pour les collectivités qui souhaiteraient les rejoindre dans cette démarche. Les collectivités peuvent également s'appuyer sur la communauté de l'Open Data. Différentes associations (LiberTIC, OpenDataFrance...) ou fondations (Fing...) contribuent aux travaux sur l'Open Data et peuvent accompagner les collectivités dans leurs démarches.

La coordination avec les acteurs locaux (autres collectivités, universités, associations, acteurs économiques, entreprises exploitant des services publics, ou entreprises liées à l'informatique et aux communications ...) est un facteur de réussite pour la démarche Open Data, car ceux-ci disposent généralement de beaucoup de données. La mutualisation des démarches Open Data entre plusieurs structures, même si elle limite la visibilité de chacune de celles-ci, permet des gains de temps et des économies d'échelle. En outre, cette mutualisation profite aux utilisateurs, qui bénéficient de jeux de données plus nombreux, plus étoffés, couvrant des périmètres géographiques et thématiques plus vastes, ce qui *in fine* favorise la création de valeur.

6.5 - Accompagner et être à l'écoute des utilisateurs

Les expériences des collectivités interviewées montrent que la simple mise à disposition des données par la collectivité n'est pas un gage de succès pour une démarche Open Data. Comme le souligne un développeur interviewé dans le cadre de l'étude, « *il ne faut pas juste mettre des données en ligne, pour communiquer et jouer la transparence, mais réfléchir à comment et pour quoi les données vont ou pourraient être réutilisées* ».

Pour faire de l'Open Data une vraie avancée pour tous, il est nécessaire de multiplier les lieux d'échanges entre secteur public, secteur privé et citoyens. En ce sens, certaines collectivités interviewées animent des journées d'échanges ou d'information, organisent des « *hackathons* »⁶² autour de la démarche Open Data. Par ailleurs, plusieurs collectivités encouragent la réutilisation des données publiques ouvertes par des concours ou des appels à projets. Il est également utile d'établir un forum de discussion permettant aux utilisateurs de confronter leurs expériences ou difficultés, d'exprimer leurs besoins, de proposer la publication de nouveaux jeux de données... Néanmoins ces événements sont destinés aux programmeurs et ne touchent pas l'ensemble des citoyens. On voit ainsi apparaître des Infolab, lieu d'expression des besoins des citoyens et lieu d'information sur le numérique.

62 Événement où des développeurs se réunissent pour faire de la programmation informatique collaborative.

Conclusion

Depuis les années 2000, le nombre de plates-formes Open Data mises en œuvre par les collectivités françaises ne cesse d'augmenter. Si elles nécessitent une organisation et des moyens spécifiques, elles modernisent les services publics, rendent plus transparente l'action publique et permettent l'apparition de nouveaux services.

L'État, qui a donné l'impulsion d'une infrastructure Open Data nationale via la plate-forme www.data.gouv.fr, suit avec intérêt les initiatives portées par les collectivités et réciproquement (Etalab est membre associé de l'association OpenDataFrance, qui fait elle-même partie du groupe des experts d'Etalab). Bien que leur collaboration avec l'État cherche son modèle, les collectivités interviewées expriment d'ores et déjà deux attentes vis-à-vis de celui-ci : l'harmonisation des bases de données à l'échelle nationale et l'extension de sa politique d'ouverture des données.

La poursuite de la politique d'ouverture des données de l'État est une attente forte de la part des collectivités et de la communauté Open Data. C'est un des axes de travail du Plan d'action de la France⁶³ pour la mise en application de la charte du G8 sur l'ouverture des données publiques⁶⁴. Le plan d'action prévoit notamment que l'État rende accessible les données à fort potentiel de réutilisation et étudie l'élargissement de la stratégie d'ouverture des données publiques aux établissements publics, aux délégataires de service public, aux autorités administratives indépendantes et aux missions de service public à caractère industriel et commercial. Dans cette perspective, les collectivités attendent également de l'État un appui pour agir auprès des grandes institutions publiques ou privées, ainsi qu'une coordination nationale des différentes initiatives autour des démarches Open Data.

La standardisation des formats de données parmi les plates-formes Open Data est un autre enjeu important des prochaines années, afin de faciliter la réutilisation des données à l'échelle nationale, voire supérieure. Le cas des données d'information routière est significatif du manque actuel d'harmonisation : les données ouvertes par les collectivités sont présentées dans des formats spécifiques, et ne sont pas conformes aux normes internationales en vigueur dans le domaine de l'information routière. Il n'est donc pas possible d'exploiter simplement les données diffusées par différents fournisseurs.

Certains acteurs de la communauté Open Data proposent des standards pour l'harmonisation des formats et référentiels au niveau national voire européen. À défaut d'une mutualisation des démarches Open Data, certains estiment que l'État devrait se saisir de ces standards et accompagner les collectivités locales souhaitant lancer ou améliorer leur plate-forme, afin de rechercher la meilleure synergie possible entre les différentes démarches et de garantir l'homogénéité des données ouvertes.

Une grande quantité de services développés par des tiers à partir des données publiques ouvertes utilisent des données de mobilité. Il est donc important pour la collectivité de bien choisir ces données, leur format et leur contenu. Le panorama des données de mobilité ouvertes montre une grande variété.

La plupart des plates-formes diffusent a minima des données géographiques sur les arrêts et les lignes.

63 Téléchargeable à l'adresse www.gouvernement.fr/gouvernement/plan-d-action-g8-sur-l-ouverture-des-donnees-publiques

64 Cette charte a été adoptée le 18 juin 2013 lors du 39ème sommet des pays du G8 (Allemagne, Canada, États-Unis, France, Italie, Japon, Royaume-Uni et Russie)

Celles-ci gagneraient à disposer d'un format commun, dans la lignée du modèle d'arrêt partagé proposé par l'AFIMB. Pour les transports collectifs, les données décrivant l'offre complète théorique sont généralement diffusées au format standard GTFS, ce qui permet à des applications pour smartphone comme *Moovit* ou *UrbanPulse* de proposer un service de recherche d'itinéraire en transport collectif sur plusieurs villes. En revanche, pour les horaires en temps réel, les formats utilisés sont variables d'une collectivité à l'autre. Des efforts de communication et d'adaptation au contexte de l'Open Data permettraient à la norme SIRI d'être plus utilisée, et permettrait une harmonisation des formats utilisés sur les plates-formes Open Data des collectivités.

En dehors des transports collectifs, un autre type de données gagnerait à être diffusé et standardisé : l'information sur le stationnement. L'exemple nantais montre le potentiel de ces données, dont la plupart des collectivités disposent déjà. Leur standardisation ne nécessite pas forcément une norme en tant que telle : un format standard promu par un collectif légitime d'acteurs du stationnement suffirait.

Une fois l'ouverture des données généralisée, dans des formats standardisés, l'accent devra alors être mis sur la qualité des données. Ce n'est que lorsque ces trois conditions seront réunies que des services innovants pourront pleinement se développer, répondant à un réel besoin des usagers, et avec un modèle économique qui pourra être pérenne. Ce chemin peut sembler long, mais le paysage est ouvert...

Annexes

Annexe A - La grille d'enquête

La grille des questions proposées lors des entretiens avec les acteurs locaux se compose de 5 parties :

A. Présentation de la plate-forme de partage de données

1. Les données et leur « périmètre »:

- a) Pourquoi tel type de données sont mises en partage et pas d'autres ?
- b) Quels types de données voudriez-vous mettre en partage et quels sont les obstacles aujourd'hui ?

2. Les modalités techniques

- a) S'agit-il de données brutes ou y a-t-il eu un formatage avant la mise en partage ?
- b) Quel est le type de format des données ? Est-ce que l'utilisateur du site dispose de précisions sur le format des données pour en faciliter la réutilisation ?
- c) Avez-vous mis en place une vérification systématique de la conformité des données publiées aux normes existantes (ex : Neptune, Siri) ?
- d) Quelle licence est utilisée (licence ouverte Etalab, ODbL, autres) ? Quelles étaient les raisons du choix de telle ou telle licence ? Quelles problématiques se posaient (attentes, difficultés...) ?
- e) Quelle réflexion avez-vous eu sur la qualité des données (certification, sélection des seules données de qualité, pas de tri des données...) ? Est-ce que des exigences sur la qualité des données ont-elles été définies ou imposées avant la mise à disposition des données ?
- f) Est-ce que la vérification de la qualité des données est faite systématiquement ? Si oui, quelle méthode (manuelle ou automatique) est utilisée ?
- g) Existe-t-il un dispositif de remontée d'erreur ? Quelles problématiques se posaient (attentes, difficultés...) ?
- h) Est-ce que les données sont simplement référencées ou hébergées ?
- i) Quels sont les relations avec d'autres démarches Open Data de collectivités sur le même territoire (commune, échange de données mutuel, autre collaboration...) ?
- j) Est-ce que les données sont internes à une collectivité, en provenance d'un opérateur de transport, de SIM ? La mise en place de convention partenariale a-t-elle été nécessaire ?
- k) Est-ce que des exigences de qualité ont été définies pour les résultats ou services issus de la réutilisation des données ?
- l) Est-ce que des exigences ont été définies sur le respect de conditions de réutilisation liées à l'intérêt général et aux politiques publiques ?

3. Le périmètre / le partenariat

- a) Quel service (transport-mobilité, géomatique, informatique...?) coordonne le projet Open Data ? Quels services ou compétences sont associés ou mobilisés dans la démarche ?

(Plus précisément, pour chaque type de données, quel service de la collectivité a été mobilisé.)

b) Quels partenaires extérieurs (en France comme à l'étranger) ont été associés à la démarche?

c) Avez-vous participé à des groupes de travail sur l'Open Data ?

4. Les moyens

a) Quel est le coût estimé de la mise en œuvre et du fonctionnement de la plate-forme ?

b) Combien d'ETP ont été dédiés au montage de la plate-forme et pour son fonctionnement / développement ?

B. Les origines du projet

a) À quelle époque vous êtes-vous lancés dans cette démarche de partage des données ?

b) Quelle est la date d'ouverture de la plate-forme de partage ? Quelle est la durée pour le montage de la plate-forme ?

c) Quel a été l'élément déclencheur de la mise en œuvre de la plate-forme ?

d) Quelles étaient vos attentes initiales ?

e) Sur quel domaine (transports, environnement...) a porté la volonté initiale de libération des données ? Pourquoi ce domaine a-t-il été ciblé selon vous ?

f) Vous êtes-vous inspiré de la démarche d'une autre collectivité (française ou étrangère) ? Si oui, laquelle ? Les avez-vous consultées ? Sur quels points ont porté votre discussion, votre observation ?

g) En ce qui concerne les données de mobilité, quels étaient vos objectifs ?

h) Quelles types de réutilisations des données de mobilité sont attendues (applications smartphones, autres...) ? Quelles en sont les motivations ?

(préciser le cas des données routières si elles sont mises à dispositions)

h) Est-ce que la réutilisation des données est libre ou encadrée (réutilisation limitée, payante...) ?

i) Quelles ont été les grandes étapes du montage du projet ? Quel est le portage au sein de la collectivité ? Quels sont partenaires ?

C. Les freins au montage et au développement de la démarche

a) Afin d'identifier les freins et attentes sur la libération et l'usage des données, avez-vous associé les exploitants ou autres détenteurs de données externes, les citoyens, les associations d'utilisateurs, ou de potentiels réutilisateurs (développeurs, universitaires, entreprises...) ?

Pour les questions suivantes, nous nous intéressons aux différents types de freins aussi bien au montage de la démarche, au fonctionnement ou au développement de la plate-forme

Open Data.

b) Quels ont été les principaux freins pour votre projet ? Quels sont, selon vous, les principaux freins pour les démarches Open Data en général ?

c) Y a-t-il eu des freins juridiques ? Sont-ils plus nombreux ou importants pour les données de mobilité ?

d) Quels sont les freins techniques ? Et particulièrement, pour les données mobilité ?

e) Quels sont les freins rencontrés en interne au sein de la collectivité ? Quels sont ceux spécifiques aux données mobilité ?

f) Quels sont les freins rencontrés avec les différents partenaires du projet ? Quels sont ceux spécifiques aux données mobilité ?

g) Est-ce que la démarche Open Data s'est faite naturellement dans la relation de négociations entre les autorités organisatrices et les opérateurs de transport ?

Est-elle contractualisée ou fait-elle l'objet d'avenants aux contrats courants ?

Est-elle prévue en amont dès l'origine ou dans le cahier des charges de la DSP ou dans d'autres documents contractuels ?

g) Qu'est-ce qui vous a manqué pour monter ce projet, de quel outil ou appui auriez-vous aimé disposer ? Un guide, un appui de l'État...

h) Plus généralement, quelle est la nature des négociations avec les opérateurs de TC ou des autres modes de déplacements (notamment VLS) ?

i) Qu'est-ce qui vous a manqué pour monter ce projet ? De quel outil ou de quel appui auriez-vous aimé disposer ?

D. Les résultats

a) Est-ce qu'une évaluation de la démarche a été réalisée ou est prévue ?

b) De manière générale, quel premier bilan faites-vous aujourd'hui de votre démarche Open Data ? Plus précisément, qu'en est-t-il pour le partage des données mobilité ?

c) Quels sont les apports de la plate-forme Open Data pour votre collectivité locale ? Y a-t-il un « plus » en ce qui concerne le partage des données mobilité ?

d) Quelles réutilisations des données ont été mises en place à ce jour ? Sur quels types de données ? Y a-t-il des réutilisations des données sur la route et quels sont les apports de ces réutilisations ?

e) Quels sont apports de la plate-forme Open Data pour le citoyen et pour le tissu économique ? Y a-t-il un « plus » en ce qui concerne le partage des données mobilité ?

f) Avez-vous lancé des concours (ou d'autres manifestations/incitations) pour impulser la réutilisation ou le développement d'outil à partir des données libérées ? Si oui, quels en sont les résultats ? Quel usage faites-vous des outils ou applications lauréates ?

E. Les améliorations à apporter / les perspectives d'évolution

a) De manière générale, quelles sont les principales pistes d'évolution pour améliorer votre de partage des données ?

- b) Quelles sont les améliorations à apporter sur le plan technique ?
- c) Quelles sont les améliorations à apporter sur le plan de la réutilisation des données ? Faut-il davantage encadrer ou libéraliser cette réutilisation ?
- d) Quel est le type de formalisation des relations ou de contrat sont envisagés avec les développeurs ? Par quel type de contrat ou avec quels objectifs en termes de services à mettre en place prioritairement passerait plus d'encadrement (si souhaité) ?
- e) Quel accompagnement à la réutilisation (pour aider les développeurs) pourrait être proposé ? (espace dédié aux développeurs, forum...)
- f) Avez-vous mis ou prévu de mettre en place un processus d'identification des réutilisateurs ? Pourquoi et comment ? Quelle connaissance avez-vous de ces réutilisateurs ?
- g) Avez-vous envisagé des actions en matière de participation citoyenne ? Quelles améliorations à apporter en termes de communication/sensibilisation/visualisation des données pour que tout le monde puisse comprendre les données en partage et les réutiliser (et pas uniquement les spécialistes) ? Quelles aides au développement des initiatives ?
- h) Quels sont vos attentes ou besoins pour faire vivre ou faire évoluer votre plate-forme Open Data ? (Notamment, quelles attentes vis-à-vis de l'État ?)
- i) Connaissez-vous/utilisez-vous Passim et Passim+ (portails annuaires des sites et services sur la mobilité et notamment des sites web Open Data) ?
Connaissez-vous Chouette (outil permettant d'échanger des données selon différents formats et de tester la conformité à la norme Neptune) ?
Êtes-vous intéressés par ce type d'outils (mis en place par le ministère du Développement durable et des Transports) ?
Comment améliorer ce type d'outils pour qu'ils répondent mieux à la demande des collectivités ?
Avez-vous besoin d'une formation pour mieux utiliser les outils à votre disposition ?
- j) Quels conseils donneriez-vous à une collectivité voulant se lancer dans l'Open Data ?

Annexe B - Liste des collectivités et entreprises interviewées

L'enquête a été réalisée durant le premier semestre 2013 auprès de 9 collectivités locales (en gras dans le tableau n° 13 ci-dessous), qui représentent globalement les différents niveaux institutionnels mettant en œuvre des plates-formes Open Data : la Région, le département, les établissements publics de coopération intercommunale (EPCI) et les communes. L'analyse des plates-formes Open Data a ensuite été élargie aux plates-formes des collectivités de niveau inférieur ou supérieur ; ainsi l'entretien avec le conseil général de la Gironde a été complété avec l'analyse des plates-formes de la Région Aquitaine, de la communauté urbaine de Bordeaux et de la Ville de Bordeaux. Le tableau n° 13 ci-dessous détaille les plates-formes analysées :

Collectivités locales	Type de collectivité	Nom de la plate-forme Open Data	Lien internet de la plate-forme Open Data	Contacts
Communauté urbaine de Bordeaux	Communauté urbaine	La CUB Open Data	data.lacub.fr	
Ville de Bordeaux	Commune	Open Data Ville de Bordeaux	opendata.bordeaux.fr	
Région Aquitaine	Région	Pigma (Plate-forme d'Information Géographique Mutualisée en Aquitaine)	pigma.org	
Département de la Gironde et Région Aquitaine	Département et Région	Ressourcerie Datalocale - portail mutualisé de données locales ouvertes	datalocale.fr	Pascal Romain : chef de projet informatique documentaire, service Projets études conseils, direction des Systèmes d'information, département de la Gironde
Grand Lyon	Communauté urbaine	Grand Lyon Smart Data	Smartdata.grandlyon.com	- Grégory Blanc-Bernard : chef de projet informatique, Grand Lyon - Grégory Bourbon : responsable du service Web et Organisation, Grand Lyon
Toulouse Métropole	Communauté urbaine	ToulouseMetropole.data [données publiques]	data.toulouse-metropole.fr	Sandrine Mathon : chef de projet Open Data et responsable du service Administration de la direction des Systèmes d'Information de la Ville et de la communauté urbaine
Montpellier Agglomération	Communauté d'agglomération	Montpellier Agglomération données ouvertes	data.montpellier-agglo.com	Pierre Brice : directeur des Ressources informatiques, direction des Ressources informatiques de l'agglomération de Montpellier
Ville de Montpellier	Commune	Territoire Montpellier Numérique - données publiques ouvertes	opendata.montpelliernumerique.fr	Jean-Marie Bourgogne : responsable du programme Montpellier Territoire Numérique, direction des Systèmes informatiques, Ville de Montpellier

Ville de Nantes et Nantes Métropole	Commune et communauté urbaine	Nantes - ouverture des données	data.nantes.fr	- Hervé Jaigu : chargé de mission Innovation, direction générale Développement économique et Attractivité internationale, direction Recherche, Innovation et Enseignement supérieur, Nantes Métropole - Gérard Le Berre, chargé de mission, administrateur de données, département général des Ressources numériques, Ville de Nantes / Nantes Métropole - Gilles Farge, directeur du service Mobilité durable, direction des Déplacements, Nantes Métropole - Alexandre Cherré, assistant de projet dans la direction Communication externe, Ville de Nantes / Nantes Métropole, en charge de l'animation du site internet / partie éditoriale, animation forum, comité utilisateur
Département de Loire-Atlantique	Département	Loire-Atlantique - ouverture des données publiques	data.loire-atlantique.fr	
Région Pays de Loire	Région	Région Pays de la Loire - ouverture des données publiques	data.paysdelaloire.fr	
Région Provence-Alpes-Côte d'Azur	Région	Open PACA	opendata.regionpaca.fr	Sophie Tendeiro : chef de projet Services numériques innovants et Mutualisation des Données publiques, service Innovation et Économie numériques, direction de l'Économie régionale, de l'Innovation et de l'Enseignement supérieur, Région PACA
Ville de Paris	Commune	ParisData - Mairie de Paris	opendata.paris.fr	Jean-Philippe Clément : chargé de mission Projets Usages et Innovations numériques, direction des Systèmes et Technologies de l'information, Ville de Paris
Rennes Métropole	Communauté d'agglomération	Rennes Métropole en accès libre	data.rennes-metropole.fr	Bernadette Kessler : responsable du service Innovation numérique, Ville de Rennes et Rennes Métropole

Tableau 13 : Liste des plates-formes des collectivités locales analysées

En se limitant à la dimension de la mobilité et des transports, les entretiens et les analyses des plates-formes Open Data nous ont conduit à pointer également des plates-formes de données ouvertes spécifiques à ces thématiques que les collectivités locales peuvent référencer sur leur propre plate-forme. Le tableau n° 14 ci-dessous (non exhaustif) décrit les principales plates-formes dédiées aux transports et à la mobilité :

Plates-formes dédiées aux transports et à la mobilité	Opérateurs	Liens internet de la plate-forme Open Data	Collectivités locales concernées (données référencées sur leur site Open Data, ou données existantes)
SNCF Open Data (Gares, TER, Intercités...)	SNCF	data.sncf.com	les Régions
Keolis-rennes.com - Les données transports de Rennes Métropole en accès libre	Keolis	data.keolis-rennes.com	Rennes Métropole
RATP Open Data	RATP	data.ratp.fr	Île-de-France
plate-forme des vélos en libre service de JCDecaux	JCDecaux	https://developer.jcdecaux.com	Amiens, Besançon, Cergy-Pontoise, Créteil, Lyon, Marseille, Mulhouse, Nancy, Nantes, Paris, Rouen, Toulouse

Tableau 14 : Plates-formes Open Data d'opérateurs de transport

Par ailleurs, des entreprises, des partenaires de l'Open Data et des développeurs ont également été interviewés. Le tableau n° 15 ci-dessous en donne la liste :

Noms	Description	Liens Internet	Contacts
Open Knowledge Foundation (OKFN)	Mouvement citoyen, né en Angleterre, qui vise à promouvoir le partage de connaissances en Europe	fr.okfn.org	Chloé Bonnet : membre du groupe Opentransport d'OKFN (Open Knowledge Foundation)
Mapize	Société spécialisée dans le développement d'outils de visualisation des données et spécifiquement dans les outils cartographiques	fr.mapize.com	François Xavier Cardi : développeur pour la société Mapize
Veni, Vidi, Libri	Association ayant pour objectif de promouvoir les licences libres ainsi que de faciliter le passage de créations sous licence libre	vvlbri.org	Benjamin Jean : juriste spécialisé dans la propriété intellectuelle ; associé d'un cabinet d'avocats ; juriste au sein de l'association Veni, Vidi, Libri (VVL) ; cocréateur de Inno Cube (société de conseils et d'expertise en innovation) ; enseignant en droit de la propriété intellectuelle à Science politique Paris et à l'université Paris 8
OpenDataFrance	Association visant à fédérer l'ensemble des acteurs engagés dans un projet de partage des données publiques en France	opendatafrance.net	Jean-Marie Bourgogne : responsable du programme Montpellier Territoire Numérique, membre d'OpenDataFrance

Tableau 15 : Autres partenaires de l'Open Data interviewés

Les analyses liées aux données portent donc sur l'ensemble des 14 plates-formes des collectivités listées ci-dessus (Tableau 13), ainsi que sur les 4 plates-formes des opérateurs (Tableau 14). Les analyses liées aux enjeux ou freins portent, quant à elles, sur les 9 collectivités interrogées (en gras dans le Tableau 13), ainsi que sur les 4 autres acteurs de l'Open Data (Tableau 15).

Annexe C - La licence ODbL et la licence ouverte de Etalab

La licence ODbL (Open Database License) est un contrat de licence de base de données favorisant la libre circulation des données. Elle est issue du projet opendatacommons.org de l'Open Knowledge Foundation. Sa traduction en français est le fruit d'une collaboration entre l'association VeniVidiLibri et la Mairie de Paris dans le cadre du projet ParisData.

Le site Internet opendatacommons.org/licenses/ODbL fournit des informations détaillées sur la licence ODbL.

Le résumé de la licence ODbL traduit et adapté au droit français est reproduite du site internet de VeniVidiLibri :

Vous êtes libres :

- de partager : copier, distribuer et utiliser la base de données ;
- de créer : produire des créations à partir de cette base de données ;
- d'adapter : modifier, transformer et construire à partir de cette base de données.

Aussi longtemps que :

- Vous mentionnez la paternité : vous devez mentionner la source de la base de données pour toute utilisation publique de la base de données, ou pour toute création produite à partir de la base de données, de la manière indiquée dans l'ODbL. Pour toute utilisation ou redistribution de la base de données, ou création produite à partir de cette base de données, vous devez clairement mentionner aux tiers la licence de la base de données et garder intacte toute mention légale sur la base de données originale.
- Vous partagez aux conditions identiques : si vous utilisez publiquement une version adaptée de cette base de données, ou que vous produisiez une création à partir d'une base de données adaptée, vous devez aussi offrir cette base de données adaptée selon les termes de la licence ODbL.
- Gardez ouvert : si vous redistribuez la base de données, ou une version modifiée de celle-ci, alors vous ne pouvez utiliser de mesure technique restreignant la création que si vous distribuez aussi une version sans ces restrictions.

Avertissement

Le résumé explicatif n'est pas un contrat, mais simplement une source pratique pour faciliter la compréhension de la version complète de la licence ODbL 1.0. Il exprime en termes courants les principales notions juridiques du contrat. Ce résumé explicatif n'a pas de valeur juridique, son contenu n'apparaît pas sous cette forme dans le contrat. Seul le texte complet du contrat de licence fait loi.

Tableau 16 : Résumé de la licence ODbL (source : <http://vlibri.org/fr/licence/ODbL/10/fr>)

La seconde licence largement utilisée en Open Data est la licence Etalab. C'est également un contrat de licence de base de données favorisant la libre circulation des données. Elle est produite par Etalab, service du Premier ministre chargé de l'ouverture des données publiques et du développement de la plate-forme française Open Data (www.etalab.gouv.fr). Etalab a réalisé la licence ouverte pour faciliter la réutilisation libre et gratuite de ces informations publiques, telles que définies par l'article 10 de la loi n° 78-753 du 17 juillet 1978.

On trouve des informations détaillées sur la licence ouverte d'Etalab sur les sites internet suivants : www.etalab.gouv.fr/pages/licence-ouverte-open-licence-5899923.html ou www.data.gouv.fr/Licence-Ouverte-Open-Licence

En résumé, cette licence stipule :

Vous pouvez réutiliser « l'Information » rendue disponible par le « Producteur » dans les libertés et les conditions prévues par la présente licence.

La réutilisation de l'Information diffusée sous cette licence

Le « Producteur » garantit au « Réutilisateur » le droit personnel, non exclusif et gratuit, de réutilisation de « l'Information » soumise à la présente licence, dans le monde entier et pour une durée illimitée, dans les libertés et les conditions exprimées ci-dessous.

Vous êtes libre de réutiliser « l'Information » :

- reproduire, copier, publier et transmettre « l'Information » ;
- diffuser et redistribuer « l'Information » ;
- adapter, modifier, extraire et transformer à partir de « l'Information », notamment pour créer des « Informations dérivées » ;
- exploiter « l'Information » à titre commercial, par exemple en la combinant avec d'autres « Informations », ou en l'incluant dans votre propre produit ou application.

Sous réserve de :

- mentionner la paternité de « l'Information » : sa source (a minima le nom du « Producteur ») et la date de sa dernière mise à jour.

Le « Réutilisateur » peut notamment s'acquitter de cette condition en indiquant un ou des liens hypertextes (URL) renvoyant vers « l'Information » et assurant une mention effective de sa paternité.

Cette mention de paternité ne doit ni conférer un caractère officiel à la réutilisation de « l'Information », ni suggérer une quelconque reconnaissance ou caution par le « Producteur », ou par toute autre entité publique, du « Réutilisateur » ou de sa réutilisation.

Tableau 17 : Résumé de la licence ouverte d'Etalab (extrait de www.etalab.gouv.fr/pages/licence-ouverte-open-licence-5899923.html)

Annexe D - Bibliographie

Agence française pour l'Information multimodale et la Billettique (AFIMB) 2013. *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.

http://www.developpement-durable.gouv.fr/IMG/130408_OpenData_Rapport%20final.pdf

Bluenove 2011. *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.

<http://www.bluenove.com/publications/actualite/publication-du-livre-blanc-open-data-quels-enjeux-et-opportunités-pour-lentreprise/>

Cete de Lyon 2013. *Open Data : données de circulation routière*, octobre 2013

Cette note technique porte sur la thématique des données ouvertes de la circulation routière. Elle présente notamment un état des lieux des pratiques d'échanges de ces données entre gestionnaires et acteurs impliqués, et des jeux de données disponibles sur différentes plateformes. Elle propose enfin une synthèse des enjeux potentiels d'un passage à l'Open Data pour les données de la circulation routière.

Cete Nord-Picardie 2013. *Open Data : synthèse bibliographique*, novembre 2013

Cette note technique propose une synthèse de la bibliographie consacrée aux démarches Open Data en France et en Europe.

Chignard S. 2012. *Open Data, comprendre l'ouverture des données publiques*, Fyp éditions, 2012.

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 ?

Commission européenne 2011. *Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des Régions. L'ouverture des données publiques: un moteur pour l'innovation, la croissance et une gouvernance transparente*, 12 décembre 2012.

Cette communication, qui s'inscrit dans le cadre de la stratégie numérique pour l'Europe, expose un ensemble de mesures visant à éliminer les obstacles existants et la fragmentation du marché des données publiques dans l'UE.

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:FR:PDF>

Conseil national du numérique 2012. *Avis n° 12 relatif à l'ouverture des données publiques*, 2012
Cet avis 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...)

http://www.cnnumerique.fr/wp-content/uploads/2012/06/2012-06-05_AvisCNNum_12_OpenData.pdf

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

Cet ouvrage très complet qui revient sur l'historique de l'Open Data, met en lumière les problématiques actuelles et donne quelques pistes d'évolution.

http://www.deciderensemble.com/images/stories/etudes/open_data/Ouverture_donnees_participation_etude_final.pdf

École des Ponts ParisTech 2011. *Pour une politique ambitieuse des données publiques : les données publiques au service de l'innovation et de la transparence*. Rapport à la Délégation aux usages de l'Internet, remis au ministre de l'Industrie, de l'Énergie et de l'Économie numérique le 13 juillet 2011.

Ce rapport présente une étude sur la réutilisation des données publiques, menée pour la Délégation aux usages de l'Internet du ministère de l'Enseignement supérieur et de la Recherche. Il met en perspective la problématique et les enjeux de l'Open Data, propose un état des lieux de la réutilisation des données publiques en France, et dessine trois scénarios prospectifs pour l'évolution future de ce mouvement.

<http://www.ladocumentationfrancaise.fr/rapports-publics/114000407/>

Ecoter 2013. Lettre hebdomadaire des TIC n° 351, avril 2013

Cette lettre évoque 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.

http://www.ecoter.org/emailing/Lettre_Hebdomadaire_Ecoter_351_12.04.13_18.04.13.pdf

Etalab 2013a. *Vade-mecum sur l'ouverture et le partage des données publiques*, septembre 2013

Ce guide vise à faciliter l'engagement des ministères dans la démarche d'ouverture et de référencement de leurs données publiques sur la plate-forme nationale www.data.gouv.fr. Il rappelle les enjeux du partage des données publiques et fournit un guide de sa mise en œuvre, recensant et répondant aux questions pratiques relatives à l'Open Data.

<http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/vademecum-ouverture.pdf>

Etalab 2013b. *Plan d'action pour la France - Charte du G8 pour l'ouverture des données publiques*, 6 novembre 2013

Ce document rappelle le contexte de la politique du gouvernement français en matière d'ouverture des données publiques, détaille les principales difficultés et les défis auxquels elle fait face, et présente les axes de travail pour l'avenir que s'est fixé le gouvernement.

http://www.gouvernement.fr/sites/default/files/fichiers_joints/plan-action-france_version_francaise.pdf

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

Ce rapport consacre un chapitre à 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.

<http://doc.openfing.org/RDPU/GuidePratiqueDonneesPubliquesv1beta.pdf>

Guillemin C. 2012. « Open Data : les transports publics libèrent leurs données », *Transport public* n° 1125, pp. 60-69, juin 2012

Cet article dresse un état des lieux de l'ouverture des données de transport par les autorités organisatrices et les principaux opérateurs de transport public en France.

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

Ce rapport (référéncé dans le rapport 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 S. 2012. *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 interne sans coproduction avec la société civile.

<http://wagner.nyu.edu/files/rudincenter/opendata.pdf>

Lecko 2012. *L'Open Data au cœur de la nouvelle relation entre collectivités, entreprises et citoyens*, 12 avril 2012

Qu'est-ce que l'Open Data? Quels bénéfices les collectivités, les citoyens et les entreprises peuvent-ils espérer en retirer ? Cette étude aborde ces questions à partir de l'analyse d'une dizaine de démarches Open Data en France, en Europe et en Amérique du Nord. Cette étude s'adresse en premier lieu aux décideurs, élus et membres des directions des services à la population qui souhaitent y voir plus clair et disposer de conseils pratiques dans l'optique de se lancer dans un projet Open Data

<http://lecko.fr/etude-open-data.html>

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

Ce rapport (référéncé dans le rapport 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 facilite donc la constitution et le partage de bases de données.

http://www.mckinsey.com/~media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.pdf

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

À 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 sur les prochaines années.

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

Cette étude analyse les enjeux liés à l'ouverture des données publiques, et propose une analyse du marché de l'ouverture et de la réutilisation de ces données. Elle donne notamment des estimations de coûts pour le montage et l'exploitation d'une plate-forme Open Data. Elle présente en outre une dizaine de retours d'expériences d'administrations ou de collectivités françaises ayant lancé une démarche Open Data.

Vickery G. 2011. *Review of recent studies on PSI re-use and related market developments*, 23 septembre 2011

Le « rapport Vickery », commandé par la Commission européenne, estime les retombées économiques potentielles de l'ouverture et de la réutilisation des données publiques. Les chiffres avancés sont très contestés par la communauté de l'Open Data.

http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/final_version_study_psi.docx

Annexe E - Glossaire

AFIMB	Agence française pour l'information multimodale et la billettique
AOT	Autorité organisatrice des transports
AOTU	Autorité organisatrice des transports urbains
API	Interface de programmation informatique d'une application
CADA	Commission d'accès aux documents administratifs
Cete	Centre d'études techniques de l'équipement
CNIL	Commission nationale de l'informatique et des libertés
CSV	Comma-separated values (valeurs séparées par des virgules)
CUB	Communauté urbaine de Bordeaux
DSP	Délégation de service public
EPCI	Établissement public de coopération intercommunale
GART	Groupement des autorités responsables de transports
GML	Generalized Markup Language (langage de balisage généralisé)
GTFS	General Transit Feed Specification (spécifications pour les flux de transports en commun)
JSON	JavaScript Object Notation (notation des objets du langage JavaScript)
KML / KMZ	Keyhole Markup Language (langage à base de balises géolocales)
ODbL	Open Database License
Open Data	Donnée ouverte
PACA	Provence-Alpes-Côte d'Azur
PSI	Public Sector Information
SGMAP	Secrétariat général pour la modernisation de l'action publique
SHP	Shapefile (format standard sous logiciel de SIG)
SIG	Système d'Information géographique
XML	Extensible Markup Language (langage de balisage extensible)

Table des matières

Sommaire.....	5
Introduction.....	6
1 -L'Open Data au sein des collectivités françaises.....	7
1.1 -Rappel des définitions.....	7
1.2 -Rappel de l'historique législatif.....	8
1.3 -Où en est-on aujourd'hui en France ?.....	10
2 -Vers l'émergence d'une nouvelle gouvernance locale et d'un modèle économique.....	12
2.1 -Un enjeu politique et démocratique fort.....	12
2.2 -Une création de valeur économique pour les territoires.....	14
2.2.1 -L'économie de la donnée.....	14
2.2.2 -Le marché de la donnée : des modèles économiques territoriaux encore fragiles.....	14
2.3 -Des retombées positives pour la collectivité.....	15
2.3.1 -Mise à disposition des données : gratuité ou redevance ?.....	15
2.3.2 -Des gains certains pour la collectivité : une modernisation des services publics.....	16
2.4 -Les données de mobilité : un levier pour répondre aux exigences des usagers des transports.....	18
3 -Comment surmonter les difficultés rencontrées ?.....	20
3.1 -Des difficultés internes à la collectivité : convaincre de la finalité de l'Open Data.....	20
3.1.1 -De faux freins internes à la collectivité.....	20
3.1.2 -... surmontables par le dialogue.....	21
3.2 -Des difficultés juridiques : existence d'un cadre juridique.....	23
3.2.1 -La protection juridique des bases de données.....	23
3.2.2 -Le choix des licences.....	24
3.2.3 -Un véritable frein : le code des marchés publics.....	28
3.3 -Des difficultés partenariales : les jeux d'acteurs.....	29
3.3.1 -Mutualisation ou stratégie d'acteurs entre les collectivités potentiellement actrices d'une démarche commune d'Open Data ?.....	29
3.3.2 -Des logiques économiques et politiques différenciées entre les collectivités et les fournisseurs de données extérieurs.....	30
3.3.3 -Le poids du lien contractuel entre les collectivités et les fournisseurs de données.....	31
3.4 -Les difficultés techniques : tout tourne autour de la donnée.....	32
3.4.1 -Le choix des formats, frein à l'usage des données ?.....	32
3.4.2 -Actualisation des données : des processus à mettre en place.....	33
3.4.3 -Les capacités d'hébergement de l'information : du temps réel à l'historisation et l'archivage.....	34
3.4.4 -Des difficultés à évaluer la qualité de la donnée.....	34
4 -Les clés du montage d'une plate-forme Open Data.....	36
4.1 -S'entourer et être soutenu.....	36

4.2 -Définir ses attentes pour définir son projet.....	37
4.3 -Être pragmatique et progressif dans la mise en œuvre.....	39
4.4 -Disposer des moyens humains et financiers nécessaires.....	39
4.5 -Ouvrir rapidement des données sur la mobilité.....	41
5 -Mobilité : le choix des jeux de données.....	44
5.1 -Les transports collectifs urbains et non urbains.....	44
5.1.1 -Principaux jeux de données sur les transports collectifs.....	44
5.1.2 -Les données de transports collectifs en temps réel.....	48
5.2 -Le vélo et les pistes cyclables.....	49
5.3 -Le stationnement (hors vélo).....	50
5.4 -Le réseau routier.....	53
5.5 -Les données relatives aux autres modes de déplacement.....	56
5.6 -Recommandations techniques sur les données de mobilité.....	56
6 -Recommandations pour la mise en œuvre d'une plate-forme Open Data.....	58
6.1 -Commencer l'aventure : être pragmatique.....	58
6.2 -Bénéficier d'un portage politique volontaire et fort.....	59
6.3 -Convaincre et communiquer.....	59
6.4 -Profiter des réseaux des collectivités et de la communauté de l'Open Data.....	59
6.5 -Accompagner et être à l'écoute des utilisateurs.....	60
Conclusion.....	61
Annexes.....	63
Annexe A - La grille d'enquête.....	64
Annexe B - Liste des collectivités et entreprises interviewées.....	68
Annexe C - La licence ODbL et la licence ouverte de Etalab.....	71
Annexe D - Bibliographie.....	73
Annexe E - Glossaire.....	77
Table des matières.....	78

Open Data in local authorities in the light of mobility data

In France, around fifty local authorities are already engaged in an Open Data initiative. They are seeking to increase the transparency of public action, bring to light an ecosystem of innovative players, and to project a positive image of their areas.

This publication provides feedback on these schemes that aim to open up data to the community. The challenges and the benefits obtained make it possible to make recommendations for the implementation of an Open Data platform.

The publication also features a special focus on mobility data. These occupy a prime place among the wide range of open data. They are usually already present in local authority information systems, and they correspond to a strong demand from transport users who are hoping for innovative services developed by third parties. Smartphone apps are the most visible part of this. But for these services to develop and remain over time, the data must be available in large numbers in standardized formats, and be of high quality.

El Open data en la administración a la luz de los datos de movilidad

En Francia, unas cincuenta administraciones locales se han lanzado ya en una acción Open Data. Intentan así que la acción pública sea más transparente, que emerja un ecosistema de actores innovadores, dando una imagen positiva a su territorio.

La presente obra es un retorno de experiencias sobre estas acciones de apertura de los datos en la administración. Las dificultades encontradas y los beneficios obtenidos permiten obtener recomendaciones para la puesta en marcha de una plataforma Open Data.

La obra propone asimismo un enfoque particular sobre los datos de movilidad, que ocupan un puesto de primera calidad entre el amplio panel de datos abiertos. Efectivamente, en general ya están presentes en los sistemas de información de las administraciones y corresponden con una expectativa fuerte de los usuarios de transporte, que esperan así servicios innovadores desarrollados por terceros. Las aplicaciones para smartphone son su parte más visible. Pero para que estos servicios se desarrollen y se perennicen, los datos deben estar disponibles en abundancia, en formatos estandarizados y ser de calidad.

© 2015 - Cerema

Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement , créé au 1er janvier 2014 par la fusion des 8 CETE, du Certu, du Cetmef et du Sétra.

Le Cerema est un établissement public à caractère administratif (EPA), sous la tutelle conjointe du ministère de l'Écologie, du Développement durable et de l'Énergie et du ministère du Logement et de l'Égalité des territoires. Il a pour mission d'apporter un appui scientifique et technique renforcé, pour élaborer, mettre en œuvre et évaluer les politiques publiques de l'aménagement et du développement durables, auprès de tous les acteurs impliqués (État, collectivités territoriales, acteurs économiques ou associatifs, partenaires scientifiques).

Toute reproduction intégrale ou partielle, faite sans le consentement du Cerema est illicite (loi du 11 mars 1957). Cette reproduction par quelque procédé que se soit, constituerait une contrefaçon sanctionnée par les articles 425 et suivants du Code pénal.

Coordination – Maquettage : service éditions Cerema/Direction technique Territoires et ville

Dépôt légal : mars 2015

ISBN : 978-2-37180-037-3

ISSN : en cours

Illustration couverture : images libres de droit – montage Cerema/DTTV

Éditions du Cerema

Cité des mobilités

25 avenue François Mitterrand

CS 92803

69674 Bron Cedex

Bureau de vente

Cerema / Direction technique Territoires et ville

2 rue Antoine Charial

CS 33927

69426 Lyon Cedex 03 – France

Tél. 04 72 74 59 59 – Fax. 04 72 74 57 80

en téléchargement gratuit sur

www.cerema.fr

Rubrique Nos éditions

La collection « Connaissances » du Cerema

Cette collection présente l'état des connaissances à un moment donné et délivre de l'information sur un sujet, sans pour autant prétendre à l'exhaustivité. Elle offre une mise à jour des savoirs et pratiques professionnelles incluant de nouvelles approches techniques ou méthodologiques. Elle s'adresse à des professionnels souhaitant maintenir et approfondir leurs connaissances sur des domaines techniques en évolution constante. Les éléments présentés peuvent être considérés comme des préconisations, sans avoir le statut de références validées.

L'Open Data en collectivité à la lumière des données de mobilité

En France, une cinquantaine de collectivités locales se sont déjà lancées dans une démarche Open Data. Elles cherchent ainsi à rendre plus transparente l'action publique, à faire émerger un écosystème d'acteurs innovants, et à donner une image positive à leur territoire.

Le présent ouvrage est un retour d'expériences sur ces démarches d'ouverture des données en collectivité. Les difficultés rencontrées et les bienfaits obtenus permettent d'en tirer des recommandations pour la mise en œuvre d'une plate-forme Open Data.

L'ouvrage propose également un focus particulier sur les données de mobilité. Celles-ci occupent une place de premier choix parmi le large panel de données ouvertes. En effet, elles sont généralement déjà présentes dans les systèmes d'information des collectivités, et elles correspondent à une attente forte des usagers de transport, qui espèrent ainsi des services innovants développés par des tiers. Les applications pour smartphone en sont la partie la plus visible. Mais pour que ces services se développent et se pérennisent, les données doivent être disponibles en grand nombre, dans des formats standardisés, et être de qualité.

Sur le même thème

Référentiel de données de l'offre de transport public – Guide de mise en œuvre et de gestion
2013

La qualité des données géographiques
2010

Échanger des données localisées – Guide juridique
2007

Connaissance et prévention des risques - Développement des infrastructures - Énergie et climat - Gestion du patrimoine d'infrastructures
Impacts sur la santé - Mobilité et transports - Territoires durables et ressources naturelles - Ville et bâtiments durables

ISSN : 2417-9701
ISBN : 978-2-37180-037-3

