

L'Open data : un outil stratégique pour l'action publique

LAURA MANEVILLE

2017-2018
Master 1 APTER

Florence Laumière (responsable pédagogique)
Xavier Cauquil & Olivier Mas (maîtres de stage)

SOMMAIRE

1. LA DONNEE DANS LES COLLECTIVITES : UNE PREOCCUPATION ANCIENNE, QUI SE RENOUVELLE AUJOURD'HUI AVEC LE DEVELOPPEMENT DE L'OPEN DATA	p. 11
1.1. Définition du vocabulaire de la donnée	p. 11
1.2. Un écosystème de la donnée complexe, encadré par une législation qui s'étoffe au fil des ans	p. 14
1.3. En quoi la donnée sert-elle l'action publique et comment est-elle produite ? ...	p. 21
2. L'UTILISATION DE LA DONNEE A TOULOUSE METROPOLE	p. 30
2.1. Toulouse, parmi les pionniers dans l'open data	p. 30
2.2. Gestion et échanges de données : entre liberté de l'ouverture des données et respect des licences	p. 35
2.3. Retour d'expérience sur le stage de master 1 à l'Observatoire de la Métropole : faire face aux obstacles en amont de l'ouverture de la donnée	p. 40
2.4. L'open data : de nombreux bénéfices et des résultats concrets pour améliorer la vie en	ville p. 45
3. ZOOM SUR LE PHENOMENE OPEN DATA : UN OUTIL STRATEGIQUE POUR LES COLLECTIVITES	p. 54
3.1. La donnée ouverte comme bien commun, favorisant l'innovation économique et sociale	p. 55
3.2. L'open data en France et à l'étranger	p. 63
3.3. De l'utilité à la nécessité de la gestion du big data : être conscient de l'enjeu pour en éviter les pièges	p. 71

REMERCIEMENTS

Je tiens à remercier tout d'abord Toulouse Métropole et le service de l'Observatoire de la Métropole de m'avoir accueillie dans leurs locaux.

Merci à Xavier Cauquil, qui m'a fait découvrir un domaine que je ne connaissais pas, celui de la donnée, à travers les missions qu'il m'a confiées. Ce sont les échanges que nous avons eus, ainsi que l'opportunité de rencontrer des personnes d'autres services de la collectivité à travers mes missions, qui m'ont donné envie de m'intéresser à ce sujet.

Merci à mon tuteur de stage, Olivier Mas, de m'avoir guidée et formée à l'utilisation du logiciel de système d'information géographique mis à disposition par la collectivité.

Je souhaite également remercier Aurélie Lavigne pour son soutien et son aide, lorsque j'étais en difficulté pour effectuer certains travaux.

Votre bienveillance et votre rigueur m'ont permis d'acquérir des connaissances en matière d'informatique et d'urbanisme, grâce à la variété des données que j'ai eu à traiter par les tâches que j'ai effectuées.

Je tiens également à remercier Sandrine Mathon - chef du service administration à la direction des systèmes d'information de la ville de Toulouse et de Toulouse Métropole, responsable de l'Open Data de la Métropole - pour m'avoir accordé de son temps et apporté des informations essentielles à la réalisation de ce mémoire.

Enfin, merci à ma professeure Florence Laumière de m'avoir suivie et encadrée pour ce mémoire, dont le sujet s'est avéré parfois complexe, mais très enrichissant par les multiples facettes qu'il recouvre et les questionnements qu'il suscite. Merci également à Laurent Jégou pour ses pistes de lecture qui m'ont bien aidées à démarrer ce mémoire.

LEXIQUE

API : Interface d'accès applicative

Big Data : « *Le terme de Big Data est utilisé lorsque la quantité de données qu'une organisation doit gérer atteint une taille critique qui nécessite de nouvelles approches technologiques pour leur stockage, leur traitement et leur utilisation. Volume, vitesse et variété sont souvent les trois critères qui permettent de qualifier une base de données de "Big Data".* »¹

CNIL : Commission Nationale de l'Informatique et des Libertés. Autorité publique qui veille à la protection des données personnelles.

Crowdsourcing : « *Pratique qui correspond à faire appel à des internautes volontaires pour proposer ou créer des contenus, répondre aux questions d'autres visiteurs. Il s'agit ainsi d'une mutualisation des ressources et des compétences (par exemple Open Street Map).* » (1)

INSPIRE : directive européenne qui oblige à rendre accessible l'ensemble des données géographiques d'un pays.

Loi CADA : loi de la liberté d'accès aux documents administratifs.

Loi NOTRe : (Nouvelle Organisation Territoriale de la République) oblige les collectivités territoriales de plus de 3 500 habitants à rendre accessibles, en ligne, les informations concernant leur territoire.

ODbL : Open Data base Licence. Licence qui oblige à mentionner la paternité des données utilisées et au repartage du travail qui y a été effectué.

Open Data : ou donnée ouverte. Démarche qui tend à la mise à disposition de données publiques ou privées, libres d'accès et réutilisables par tout type de public.

RGPD : Règlement Général sur la Protection des Données. Cette loi apporte de nouvelles règles en matière de protection des données personnelles.

¹ OpenData Lab Occitanie. *Lexique autour de l'open data* [en ligne]. (2018) Disponible sur : <http://www.opendatalab.fr/l-opendata/lexique-autour-de-l-open-data> (consulté le 24/05/18)

INTRODUCTION

La société est aujourd'hui hyper connectée par l'intermédiaire des Smartphones, ordinateurs, GPS et autres technologies, toujours plus petites et nomades afin de suivre et d'être au plus près de chaque individu. Tous les utilisateurs transmettent des données personnelles. En retour, ils sont ciblés par diverses propositions de services, qu'elles soient payantes ou gratuites. Finalement, ces échanges de données orientent la vie des utilisateurs au quotidien que ce soit dans leur consommation, leurs itinéraires ou encore leurs usages. La donnée est partout, elle influence le quotidien de la population mais aussi celui des territoires. Il y a quelques années encore, bon nombre de collectivités locales se contentaient d'amasser la donnée, sans forcément l'exploiter. Or aujourd'hui, le développement des applications numériques aux mondes de l'aménagement du territoire, de l'urbanisme, de la gestion des réseaux et services urbains... se multiplient, en faveur du développement des territoires.

Au fil du temps, le législateur s'est attaché à encadrer l'utilisation des données produites. Peu à peu, les pouvoirs publics ont pris conscience du pouvoir associé à la maîtrise de données, mais aussi à leur ouverture car elle peut être bénéfique pour les collectivités : rapprochement de l'administration avec les citoyens, retombées économiques, relations entre les différents services et partenaires des territoires... L'Open Data est l'expression consacrée à l'ouverture des données publiques, c'est-à-dire à leur diffusion auprès de tous les citoyens, via Internet et ce gratuitement et avec la possibilité de réutiliser ces données. Ce processus sous-tend au mouvement de l'OpenGov ou « démocratie ouverte », qui se base sur trois principes : la transparence, la participation et la collaboration.

Toutefois, l'open data présente aussi des limites et des risques qu'il convient d'identifier et d'analyser, comme la marchandisation des données ou encore leur sécurité. Par ailleurs, les données générées par des éléments extérieurs aux collectivités peuvent parfois faire courir des risques aux territoires qui se voient directement impactés par des comportements changeant de la part des habitants et ce en très peu de temps. C'est pourquoi ces dernières décennies différentes lois ont été élaborées afin d'inciter les collectivités à faire preuve de transparence en libérant ces données, mais aussi de les obliger à s'emparer de cette

question par nécessité, afin d'éviter aux collectivités locales d'être dépassées par l'influence grandissante du big data sur leur territoire. Le dernier texte de loi qui fait référence dans le domaine est la loi du 7 octobre 2016 pour une République Numérique². Ce texte prévoit notamment la création d'un service public de la donnée et impose à tous les organismes publics et à toutes les collectivités de plus de 3 500 habitants et 50 agents de rendre accessibles leurs données en ligne d'ici à octobre 2018.

J'ai réalisé mon stage de Master 1 à Toulouse Métropole, dans le service de l'Observatoire de la Métropole. Ce service est rattaché à la Délégation à l'Aménagement et a en charge la publication de trois documents chaque année. Ces documents sont des recueils de données statistiques à l'échelle de la métropole, de la ville de Toulouse et de ses quartiers. Durant ce stage, j'ai assisté à la collecte des données qui servent à ces publications par mon service. C'est en observant cette collecte et en m'intéressant à des bases de données ouvertes (INSEE) et fermées (PERVAL par exemple), que je me suis questionnée sur l'importance de la donnée dans les collectivités. Comment est-elle produite ? Dans quel objectif ? A quoi sert-elle, quelles en sont les applications concrètes ? Puis je me suis renseignée sur les phénomènes de l'open data et sur la manière dont il est abordé à Toulouse et ailleurs. Par quel processus les données sont-elles libérées ? Qui a en charge la sélection de ces données ? Autant de questions qui m'ont poussée à rédiger ce mémoire sur le thème de la donnée, de son rôle et de son ouverture dans les collectivités. C'est pourquoi ce mémoire s'attache à répondre à la problématique suivante : **En quoi la donnée, et son ouverture, constitue-t-elle un outil stratégique pour les collectivités ?**

Pour répondre à cette interrogation, ce mémoire s'attachera dans un premier temps à poser les cadres lexicaux, législatifs et historiques autour de la donnée et de son ouverture, avant de questionner en quoi cela est utile à l'action publique. La seconde partie sera consacrée à Toulouse. En lien avec mon stage, nous questionnerons l'historique de l'ouverture

² MINISTERE DU NUMERIQUE. *LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique*. [en ligne]. Journal officiel, n°0235 du 08 octobre 2016. Disponible sur : <https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=3EFDF130F9E755A0B3DFC0AD340AAABB.tplgfr21s3?cidTexte=JORFTEXT000033202746&categorieLien=id> > (consulté le 04/06/18)

de la donnée dans la métropole, puis expliquerons comment la donnée sert l'action publique dans la métropole toulousaine. Cette partie montrera aussi que des obstacles existent, mais que les bénéfices à tirer de la maîtrise et de l'ouverture des données ne sont pas négligeables. Enfin, la troisième partie se concentrera sur l'open data. Elle exposera plus précisément les avantages que les collectivités peuvent en tirer et illustrera cette dimension à travers des exemples choisis en France, mais aussi à l'étranger. Enfin, un constat sera fait sur la nécessité de maîtriser la donnée, tout en questionnant les limites liées à l'ouverture des données.

1. LA DONNEE DANS LES COLLECTIVITES : UNE PREOCCUPATION ANCIENNE, QUI SE RENOUVELLE AUJOURD'HUI AVEC LE DEVELOPPEMENT DE L'OPEN DATA

1.1. Définitions du vocabulaire de la donnée

La donnée s'inscrit dans le domaine du numérique et de l'informatique. Cette notion est souvent accompagnée de termes plus ou moins clairs, issus parfois d'anglicismes comme : data, open data, big data, licences, API... Il est donc important de poser des définitions claires sur ces termes afin de bien comprendre ce qui sera exposé par la suite.

- **La donnée, « data »** en anglais, est la « représentation d'une information sous une forme conventionnelle, de façon à faciliter son traitement par un ordinateur »³
- Le **big data** correspond à des données massives et à leur traitement, qui nécessitent une capacité d'analyse qui va au-delà de celle d'outils informatiques classiques ou de la capacité humaine à les traiter.
- **Une base de données** est : « un ensemble structuré et organisé de données qui représente un système d'informations sélectionnées de telle sorte qu'elles puissent être consultées par des utilisateurs ou par des programmes »⁴. La donnée est à la base de l'information et de la connaissance. En effet, c'est en collectant et en présentant la donnée que l'on crée de l'information, et c'est à partir de l'information qu'il va être possible de faire des choix et donc de produire de la connaissance.

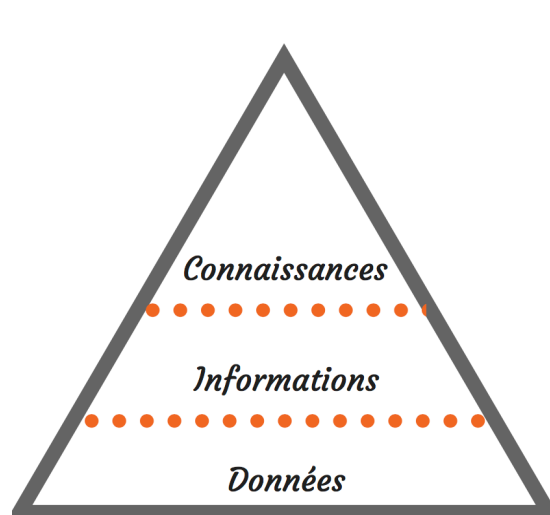


Illustration 1 : La donnée à la base de la construction de la connaissance (source : Laura Maneville)

³ GUALINO, J. *Dictionnaire pratique informatique, Internet et nouvelles technologies de l'information et de la communication*. Paris : Gualino, 2006, 507 p.

⁴ Larousse. *Base de données* [en ligne]. In : Encyclopédie Larousse. Disponible sur : http://www.larousse.fr/encyclopedie/divers/base_de_donnees/185906 > (consulté le 10/05/18)

➤ **Une licence** est un contrat entre le titulaire des droits d’auteur sur un logiciel, de la donnée ect. et un cocontractant. Elle permet de définir comment la donnée par exemple peut être utilisée, diffusée ou modifiée. Il existe différents types de licence, mais afin d’éviter la création d’une multitude de licences, qui pourraient porter à confusion, la loi pour une République numérique a dressé une liste des licences qui peuvent être utilisées par les administrations pour l’ouverture de leurs données⁵. La **licence ouverte**⁶ a été créée par la mission Etalab, dans le but de faciliter la réutilisation des données publiques. L’utilisateur est libre de :

- Partager (copier, distribuer et utiliser) la base de données ;
- Produire des créations à partir de la base de données ;
- Adapter (modifier, transformer, construire) la base de données ;
- En faire un usage commercial

Tout est possible, tant qu’il mentionne la paternité. Cette licence n’oblige pas à repartager son travail sous la même licence, contrairement à la licence ODbL. Elle favorise donc plutôt les réutilisateurs qui souhaitent faire de la donnée une réutilisation commerciale.

➤ La **licence ODbL (6)** (Open Database License) est un autre type de licence, créée par l’Open Knowledge Foundation (association anglaise qui promeut notamment l’open data). C’est la deuxième licence la plus courante en France. Comme la licence ouverte, elle permet de tout faire : copie, modification usage commercial. Tout cela à condition de mentionner la paternité et de repartager la base de données travaillée sous la même licence. Cette licence est utile pour aider à la réutilisation par de « petits acteurs » qui vont travailler de manière plus collaborative.

L’idée est de favoriser la création de biens communs, et donc l’innovation sociale, en repartageant ce qui a été produit grâce à l’utilisation d’une donnée, sous la même licence

⁵ Data.gouv.fr. *Licences de réutilisation* [en ligne]. Disponible sur : <<https://www.data.gouv.fr/fr/licences>> (consulté le 10/05/18)

⁶ CHASSAGNE, C. (dir.). *Open data - ouverture, exploitation, valorisation des données publiques*. Bourgoin Jallieu : territorial éditions, 2017, 134 p.

et donc en licence libre. Toutefois il est possible d'obtenir une dérogation pour éviter ce principe de « share alike ⁷ ».

- La **donnée publique** est produite ou collectée par l'Etat, une collectivité ou tout autre organisme public ou parapublic. Comme son nom l'indique, elle est publiée ou tenue à disposition du public.

Selon le dictionnaire Larousse, l'**open data** est : « la mise à disposition de données numériques, d'origine publique ou privée, libres d'accès et réutilisables par tous »⁸. Ces données sont disponibles en ligne et publiées de manière gratuite et sous une licence qui permet leur réutilisation, même à des fins commerciales. Conformément à la loi informatique et libertés⁹, on note que toute publication de données personnelles, par lesquelles il serait possible d'identifier un individu, est interdite. D'autre part, certaines données considérées comme sensibles peuvent ne pas être divulguées. C'est le cas des données qui peuvent mettre en jeu la sécurité nationale ou pour lesquelles des droits de propriété intellectuelle s'appliquent encore pour des tiers (exemple de données sensibles : celles qui concernent l'origine ethnique, la santé, l'orientation sexuelle des individus, ménages...).

Les données ouvertes sont publiées soit sous la forme de fichiers statistiques, soit sous la forme d'**API** (interface d'accès applicative). Les API permettent d'avoir accès à des données en temps réel, des données statistiques qui sont mises à jour instantanément ou des données trop volumineuses comme celles d'une carte géographique par exemple¹⁰.

⁷ *Share alike* : partage à l'identique

⁸ Larousse. *Open data* [en ligne]. In : Encyclopédie Larousse. Disponible sur : https://www.larousse.fr/dictionnaires/francais/open_data/10910872 (consulté le 10/05/18)

⁹ REPUBLIQUE FRANCAISE. *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés* [en ligne] Journal officiel du 7 janvier 1978 page 227. Disponible sur : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460&categorieLien=id> (consulté le 04/06/18)

¹⁰ LiberTIC. *Sensibilisation Open Data*. In : SlideShare [en ligne]. (4 avril 2017) Disponible sur : <https://fr.slideshare.net/libertic/sensibilisation-open-data> > (consulté le 12/05/18)

Le lien entre smart city et données

La smart city, ou ville intelligente, est un concept qui a pour objectif « d’optimiser ses coûts, son organisation et le bien-être des habitants »¹¹. Grâce à l’utilisation de divers capteurs, applications, objets connectés, outils numériques... la ville optimise ses flux (routiers, piétons...), devient économe en énergies, s’inscrit dans une démarche durable et tente d’associer ses habitants dans des projets participatifs voire co-construits.

Pour fonctionner, le concept de smart city se nourrit donc de la donnée, qu’il analyse, croise, enrichit, afin de créer une ville nouvelle. La data mais aussi son ouverture (open data) est un élément incontournable de la smart city, qui permet de favoriser les échanges et donc de répondre plus facilement aux besoins exprimés par les usagers de la ville. La participation des grands opérateurs d’énergies dans la mise en commun des données est également importante, afin de mieux gérer les ressources énergétiques en ville, toujours dans un objectif d’optimisation des flux.

1.2. Un écosystème de la donnée complexe, encadré par une législation qui s’étoffe au fil des ans

En 1789, l’article 15 de la Déclaration des Droits de l’Homme et du citoyen stipule que « la société est en droit de demander compte à tout agent public de son administration ». De fait, dès le XVIIIème siècle, le principe de transparence de l’administration publique (qui ne se limite d’ailleurs pas aux données) envers les citoyens est un droit fondamental. Au fil des décennies, l’informatique devient de plus en plus présente dans nos modes de vie. Plusieurs lois vont se succéder afin d’encadrer l’utilisation et la diffusion des informations numériques.

¹¹ Laurence. *C’est quoi la smart city ? Une introduction compréhensible*. In : OpenDataSoft [en ligne]. (29 avril 2016) Disponible sur : <<https://www.opendatasoft.fr/2016/04/29/cest-quoi-la-smart-city-une-introduction-a-la-ville-intelligente/>> (consulté le 23/05/18)

Dans les années 1960 l'Etat est souvent critiqué pour sa culture du secret et la population réclame plus de transparence et d'information. En 1974 le journal Le Monde publie un article sur un projet du gouvernement lancé en 1972 et nommé SAFARI¹² (Système Automatisé pour les Fichiers Administratifs et le Répertoire des Individus) ; cet article va secouer l'opinion publique. Ce projet a pour objectif de créer un fichier dans lequel chaque français se verrait identifié par son numéro de sécurité sociale, auquel serait associé un ensemble d'informations administratives le concernant. C'est ainsi qu'en 1978, afin de protéger la vie privée et les libertés individuelles, une commission est créée. Elle sera par la suite renommée la CNIL (Commission Nationale de l'Informatique et des Libertés), et veillera à ce que l'informatique reste au service du citoyen tout en respectant ses droits. L'année 1978 est aussi celle de l'adoption de la loi CADA¹³, la « loi de la liberté d'accès aux documents administratifs », qui permet le libre accès aux documents administratifs.

En 2003 la directive européenne 2003/98/CE relative à la réutilisation des informations du secteur public, incite à publier des données et facilite leur réutilisation. La directive INSPIRE en 2007 oblige les autorités publiques à rendre accessibles les données environnementales et géographiques récoltées. En France, la mission Etalab est créée en 2011 afin de « coordonner l'action des services de l'Etat et de ses établissements publics pour faciliter la réutilisation la plus large possible de leurs informations publiques »¹⁴. La même année une plateforme internationale nommée OGP (Open Government Partnership) est lancée afin de soutenir les Etats membres dans leur politique d'ouverture des données, mais aussi en matière de transparence de l'action publique et d'utilisation des nouvelles technologies pour moderniser l'action publique. Aujourd'hui, ce mouvement qui tend à l'ouverture des données est nommé l'OpenGov (« démocratie ouverte »). Il s'appuie sur trois piliers : la transparence, la

¹² ORYGA. *L'historique des lois de protection des données en France* [en ligne]. (16 janvier 2018) Disponible sur : < <https://www.oryga.com/actualites-rgpd/historique-lois-protection-des-donnees-france/> > (consulté le 04/06/18)

¹³ REPUBLIQUE FRANCAISE. *LOI n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal* [en ligne] Journal officiel, du 18 juillet 1978 page 2851. Disponible sur : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000339241&categorieLien=id> (consulté le 04/06/18)

¹⁴ Le blog d'Etalab (Etalab). *Etalab* [en ligne]. Disponible sur : <https://www.etalab.gouv.fr/qui-sommes-nous> (consulté le 09/05/18)

participation et la collaboration¹⁵. La France a rejoint l'OGP en 2014, ce qui l'oblige à élaborer un plan d'action national pour formaliser ses engagements. Ce plan a été publié en 2015 : la France a pris 26 engagements, sur 5 axes¹⁶ :

1. Transparence de l'action publique ;
2. Ouverture des ressources numériques et innovation ouverte ;
3. Démarches de participation citoyenne renforcées ;
4. Le gouvernement ouvert au service des enjeux mondiaux de notre siècle, développement, environnement et science ;
5. L'ouverture des juridictions et des autorités administratives indépendantes.

Cette année-là la France prend également la 3^{ème} place du Classement Open Data Index. Cet indicateur évalue l'ouverture des données publiques dans 184 pays au travers de dix jeux de données tels que les horaires des transports, les statistiques nationales, les résultats des élections...

En 2015 la loi NOTRe¹⁷ (Nouvelle Organisation Territoriale de la République) oblige les collectivités territoriales de plus de 3 500 habitants à rendre accessible en ligne les informations concernant leur territoire. Le 7 octobre 2016, la loi pour une République Numérique est promulguée : « elle vise à favoriser l'ouverture et la circulation des données et du savoir, à garantir un environnement numérique ouvert et respectueux de la vie privée des internautes et à faciliter l'accès des citoyens au numérique »¹⁸. Cette nouvelle loi marque la création d'un service public de la donnée, géré par la mission Etalab, qui permettra de mettre

¹⁵ Open Data Lab Occitanie. *L'open data, c'est quoi ?* [en ligne]. (2018) Disponible sur : < <http://www.opendatalab.fr/l-opendata/qu-est-ce-que-l-opendata> > (consulté le 30/03/18)

¹⁶ Secrétariat d'État à la réforme de l'Etat et à la simplification. *Pour une action publique transparente et collaborative : plan d'action national pour la France 2018-2020* [en ligne]. (2017) Disponible sur : < <https://www.etalab.gouv.fr/wp-content/uploads/2018/04/PlanOGP-FR-2018-2020-VF-FR.pdf> > (consulté le 03/05/18)

¹⁷ REPUBLIQUE FRANCAISE. *LOI n° 2015-991 du 7 août 2015 portant nouvelle organisation territoriale de la République* [en ligne] Journal officiel n°0182 du 8 août 2015 page 13705. Disponible sur : < <https://www.legifrance.gouv.fr/eli/loi/2015/8/7/2015-991/jo/texte> > (consulté le 04/06/18)

¹⁸ Gouvernement.fr. « Pour une République Numérique » [en ligne]. (15/05/17) Disponible sur : <https://www.gouvernement.fr/action/pour-une-republique-numerique> (consulté le 09/05/18)

en ligne des données importantes en matière économique ou sociale. Ces données sont destinées en premier lieu aux entreprises et aux administrations car ce sont des structures qui ont directement besoin de ces données pour travailler. A ce jour, neuf jeux de données « de référence » ont été identifiés (cf. illustration 1). Ils sont disponibles via le site data.gouv.fr, qui est géré par la mission Etalab (cf. illustration 2). Depuis la loi pour une République Numérique, les données ne peuvent plus être monnayées, comme cela pouvait être le cas auparavant.

Illustration 2 : Les 9 jeux de données « de référence » identifiés (extrait du site : data.gouv.fr)

Jeux de données à la une

Derniers jeux de données



Base Sirene des entreprises et de leurs établissements (SIREN, SIRET)

La base Sirene contenant des données à caractère personnel, l'Insee attire votre attention sur les obligations légales qui en

381 37 52 France



Cadastre

Le plan cadastral est le découpage du territoire français en unités de surfaces permettant le calcul de certains impôts.

1 2 17 France



Plan Cadastral Informatisé (PCI)

Le plan cadastral est un assemblage d'environ 600 000 feuilles ou planches représentant chacune une section ou une

3 2 48 France



BAN - Base Adresse Nationale

La Base Adresse Nationale est une base de données qui a pour but de référencer l'intégralité des adresses du territoire

3 26 31 France Point d'Intérêt



Codes postaux

Ce jeu de données fait le lien entre les codes postaux et les adresses. Il est construit à partir du Code Officiel

2 0 2 France Commune française



Délimitation Parcellaire des AOC Viticoles de l'INAO

Attention : Les données communiquées ici ne le sont qu'à titre informatif. Les délimitations parcellaires officielles sont

3 4 5 01/01/2010 à 19/03/2018 France



Répertoire des logements locatifs des bailleurs sociaux (RPLS)

Le répertoire des logements locatifs des bailleurs sociaux a pour objectif de dresser l'état global du parc de logements locatifs

3 2 2 01/01/2015 à 29/11/2017 France



Annuaire santé de la Cnam

L'annuaire santé répond à la mission générale d'information des assurés en vue notamment de faciliter l'accès aux soins,

15 1 2 Pays



Liste Publique des Organismes de Formation (L.6351-7-1 du Code du Travail)

Conformément à l'article L.6351-7-1 du Code du Travail, la liste des organismes déclarés auprès du Préfet de Région

3 0 18 France Point d'Intérêt



Découpage administratif communal français issu d'OpenStreetMap

Exports du découpage administratif français au niveau communal (contours des communes) issu d'OpenStreetMap produit

13 26 39 France Commune française



Jurisprudence anonymisée de la Cour des comptes (2006-2008 et 2010-2015)

Dans le cadre de sa #DataSession des 27 et 28 mai 2016, les juridictions financières ont décidé d'ouvrir en open data le texte

8 2 0 France



Aides à la presse : classement des titres et groupes de presse aidés

Dans un objectif de transparence, le ministère de la culture publie la liste des titres et, à partir de 2016, des principaux

11 5 5 France

Illustration 3 : page d'accueil du site data.gouv.fr



Enfin, en matière de protection des données personnelles, il faut citer la loi RGPD¹⁹, une loi européenne, qui vient d'entrer en vigueur depuis le 25 mai 2018. Le Règlement Général pour la Protection des Données (RGPD) met en place des obligations en matière de protection de données personnelles et de transparence vis-à-vis de leur utilisation : "On doit vous dire très clairement pourquoi est-ce qu'on collecte vos données, pour quelle finalité. On doit vous dire qui va être destinataire de ces données et qui y aura accès. Vous devez aussi savoir combien de temps on va conserver vos données. Et on doit vous dire quels sont les droits que vous avez par rapport à ces données" Thomas Dautieu, directeur conformité de la CNIL²⁰.

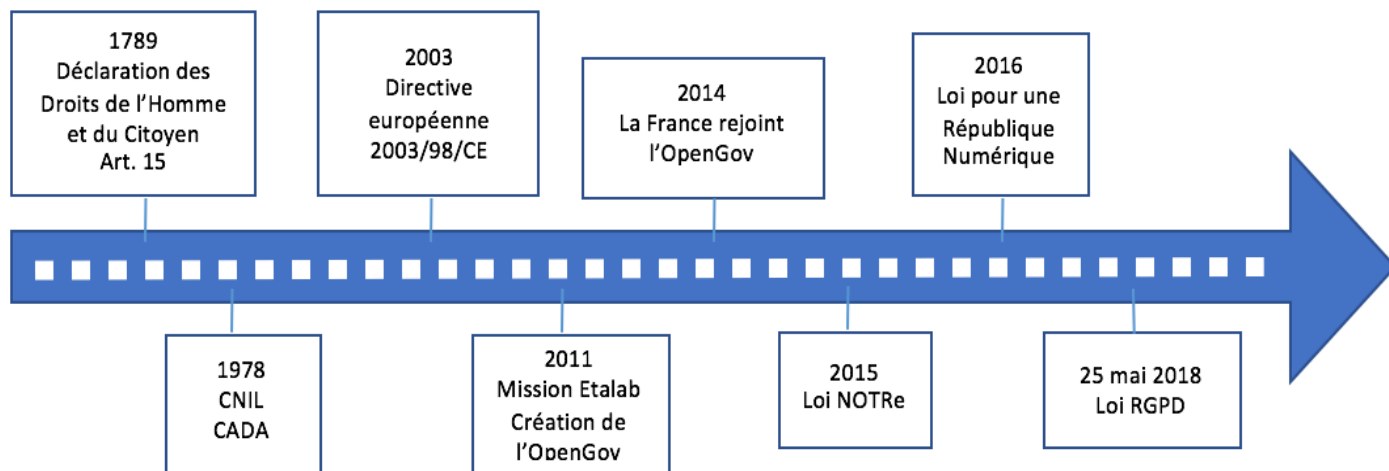
¹⁹ LEGIFRANCE. *Projet de loi relatif à la protection des données personnelles* (JUSC1732261L) [en ligne]. (2018) Disponible sur :

<https://www.legifrance.gouv.fr/affichLoiPreparation.do;jsessionid=AD5660270AD9F70B94275AC823321680.tplgfr22s_3?idDocument=JORFDOLE000036195293&type=contenu&id=2&typeLoi=proj&legislature=15 >

(consulté le 04/06/18)

²⁰ GEFFRIER F., MARTINEZ G., LEROY A. *RGPD : que va changer la nouvelle loi sur les données personnelles ?*) [en ligne]. (2 mai 2018) Disponible sur : < <http://www.europe1.fr/technologies/rgpd-que-va-changer-la-nouvelle-loi-sur-les-donnees-personnelles-3641006> > (consulté le (04/06/18)

Illustration 2 : Les grandes lois de la donnée au fil du temps (source : Laura Maneville)

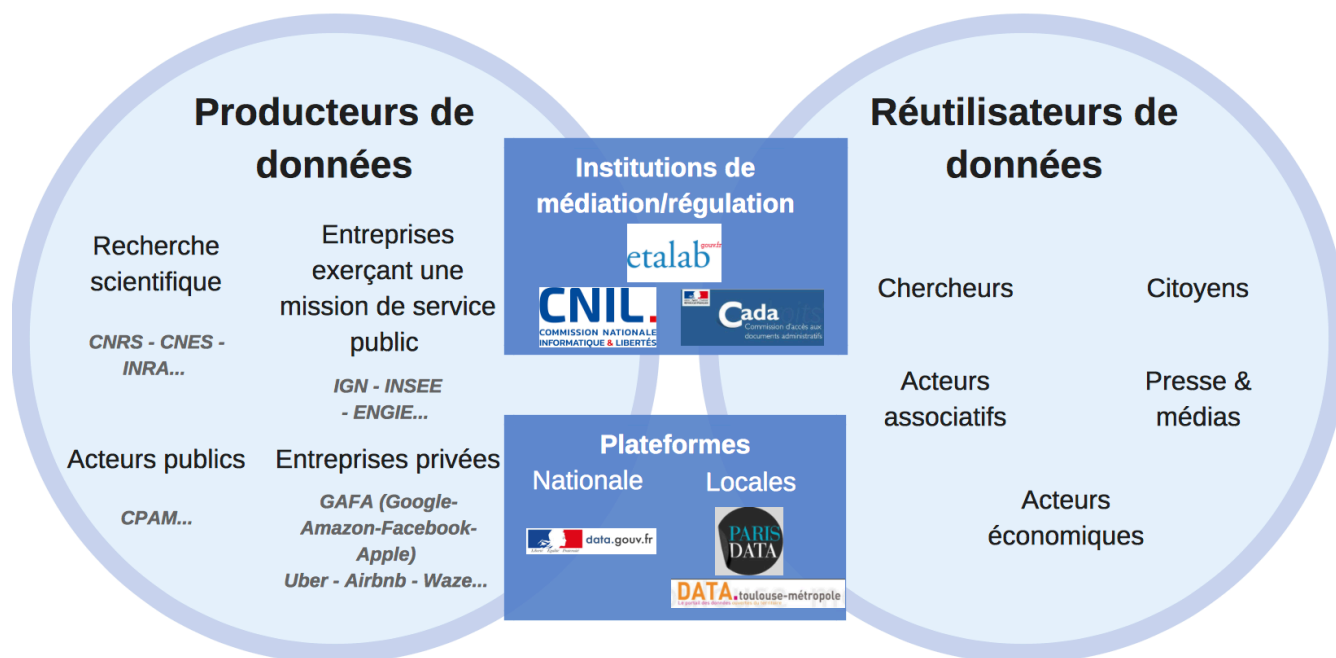


Le réseau d'acteurs qui agit autour de la donnée est en constante évolution et regroupe toutes les sphères de notre société. L'écosystème de la donnée rassemble des acteurs qui peuvent être distingués en deux groupes :

- Les producteurs de données et les réutilisateurs. Les producteurs rassemblent des instituts de recherche scientifique (CNRS, CNES, INRA...), des organismes publics (CPAM, CAF...), des entreprises privées (GAFA, Uber, Airbnb, Waze...) ou encore des entreprises exerçant une mission de service public (IGN, INSEE, ENGIE...).
- Les réutilisateurs se composent des citoyens, des acteurs économiques, de la presse et des médias, des chercheurs et des acteurs associatifs.

Les activités de ces deux groupes d'acteurs sont régulées par le législateur, aidé par des institutions comme la CNIL ou la CADA, qui est une commission d'accès aux documents administratifs, ainsi que la mission Etalab, qui coordonne l'action des acteurs publics pour faciliter la réutilisation de la donnée publique. Les données sont mises à disposition sur une plateforme nationale (data.gouv.fr) ou des plateformes locales (Open Data Toulouse Métropole, Open Data Paris...).

Ecosystème de la donnée



La donnée, un bien marchand ?

Avant la loi pour une République Numérique (2016), les données que les collectivités récupéraient auprès des entreprises privées pouvaient être payantes. La collectivité pouvait donc acheter des données à l'INSEE, à l'IGN... Cette marchandisation de la donnée a été remise en question, notamment car cela signifiait que les collectivités achetaient des données à des entreprises qui étaient subventionnées indirectement par l'Etat²¹, à travers les routes, les pistes cyclables... Par exemple, une entreprise qui propose des vélos en libre-service dans la ville profitait gratuitement des pistes cyclables offertes par la collectivité, mais pouvait vendre à une collectivité les données qu'elle produisait entre autre grâce à ces pistes cyclables. La

²¹ LeLab, CHRONOS, OUI SHARE. *Quelles coopérations public-privé à l'ère de la data ?* [en ligne]. In : Institut ville durable - Les dessous de l'innovation, n°2. (mai 2017) Disponible sur : < http://institutvilledurable.fr/wp-content/uploads/2017/10/CHRONOS-ENQUETES-n2_VF-VERSION-WEB.pdf > (consulté le 30/05/18)

collectivité se retrouvait donc cliente de l'entreprise en achetant des données qui lui étaient utiles à l'amélioration des politiques publiques.

L'introduction de la notion de « **donnée d'intérêt général** » et leur ouverture, grâce à la loi pour une République Numérique, a changé la donne. Les données d'intérêt général sont définies comme « des données qui sont de nature privée mais dont la publication peut se justifier en raison de leur intérêt pour améliorer les politiques publiques »²². Grâce à cela, les données définies comme telles sont aujourd'hui partagées gratuitement. Ainsi, des bases comme PERVAL, qui donnent des indications sur les transactions immobilières et peuvent donc servir aux politiques d'aménagement et d'urbanisme, sont aujourd'hui libres d'accès pour les collectivités, puisque d'intérêt général. Les opérateurs publics, tels que l'INSEE ou encore l'IGN, n'ont donc plus le droit de vendre leurs données aujourd'hui.

1.3. En quoi la donnée sert-elle l'action publique et comment est-elle produite ?

Depuis longtemps, l'Etat a compris l'importance de la donnée afin d'assurer le bon fonctionnement et l'organisation du pays, aussi bien au niveau économique que social. La connaissance de la population est l'une des premières données collectées, à travers les registres paroissiaux au XVII^e siècle, dans l'intérêt de calculer l'impôt²³. L'Etat a également collecté des données sur son territoire, à travers des campagnes de cartographie sous Louis XIV, afin d'assurer sa défense. Plus la donnée est collectée et analysée, plus la connaissance du territoire est fine et plus le pays peut être contrôlé, organisé et administré correctement. Encore aujourd'hui, on remarque que le ministère qui gère une donnée correspond à la finalité

²² economie.gouv.fr. *République numérique : ouverture des données d'intérêt général* [en ligne]. (22/09/2016) Disponible sur : < <https://www.economie.gouv.fr/republique-numerique-ouverture-donnees-d-interet-general> > (consulté le 01/06/18)

²³ Secrétariat général pour la modernisation de l'action publique. *Rapport au premier ministre sur la gouvernance de la donnée 2015 – la donnée au service de la transformation de l'action publique* [en ligne]. (décembre 2015) Disponible sur : < https://www.gouvernement.fr/sites/default/files/document/document/2016/01/rapport_au_premier_ministre_sur_la_gouvernance_de_la_donnee_2015.pdf > (consulté le 01/06/18)

de celle-ci : par exemple la cadastre est confié à la direction générale des finances publiques. Après la seconde guerre mondiale la statistique publique est associée à la planification, afin de combiner la connaissance de la population et de l'économie avec de grands projets d'aménagement et de développement du territoire.

A partir des années 1970 l'informatisation des administrations a permis de transformer les registres en systèmes d'information. Toutes ces informations réunies demandent de nouvelles capacités de traitement, de calcul et d'analyse. La donnée peut être produite de différentes manières, elle est de nature multiple (statistiques, informations météorologiques, budgétaires...) et ses utilisations peuvent être multiples, c'est pourquoi elle est difficile à maîtriser, à organiser et à gouverner.

L'Etat, les collectivités territoriales, les administrations publiques, produisent de la donnée. Toutefois l'Etat a choisi de créer des opérateurs consacrés à la production de données, afin d'avoir des bases fiables et très précises plus facilement en déléguant cette tâche, mais aussi actualisées plus rapidement. L'Institut National de la Statistique et des Etudes Economiques (INSEE) a en charge la collecte, l'analyse et la diffusion de statistiques sur la population, l'économie et le territoire en général. Cet opérateur s'occupe également des registres d'état civil et des répertoires d'entreprises (répertoire Sirene). L'IGN (Institut Géographique National) s'occupe de la cartographie du territoire. Un troisième opérateur à retenir serait l'INSERM (Institut National de la Santé et de la Recherche Médicale). Cet institut gère la recherche liée à la santé : un très grand nombre de données est produit concernant entre autre l'épidémiologie.

Aujourd'hui, la production de données repose également sur des entreprises privées, ou même des citoyens, et cela représente un volume considérable d'informations recueillies voire traitées. Cependant il convient de garder à l'esprit que les données produites par ces structures privées peuvent être alimentées par des sources non vérifiées et non scientifiques, comme des contributions d'internautes par exemple. Elles doivent donc être utilisées avec prudence, car leur fiabilité peut être remise en cause. Les données privées et publiques peuvent toutefois être combinées afin de produire de nouvelles connaissances. La production des données peut aussi se faire de manière collaborative. L'association OpenStreetMap est

établie au niveau international. Elle collecte des données pour créer une carte libre du monde. L'association d'OpenStreetMap France avec l'IGN, La Poste et l'Administrateur général des données a permis la création de la Base adresse nationale : elle permet de faire correspondre une adresse postale à une adresse géographique. Distribuée sous licence ODbL, elle est réutilisable par n'importe quelle structure et peut être utile à la création de multiples applications, cartes etc.

Pourquoi récolter et utiliser la donnée ?

A travers les lois sur la donnée et son utilisation, l'Etat répond à une double attente de la part de la population et des acteurs économiques : la protection des données personnelles, mais aussi l'ouverture des données dont il dispose auprès d'un large public. Mais la donnée est également bénéfique au gouvernement car elle est un outil pour mieux administrer le pays. Elle permet d'améliorer l'efficacité des services publics (Mes-aides.gouv.fr est un portail qui permet à tout citoyen de connaître l'ensemble des prestations sociales auxquelles il a droit), et aide à la construction et à l'application des politiques publiques. L'open data est aussi source, à travers la mise en commun des données, de création de nouvelles connaissances et d'innovations (grâce à la création d'applications pour améliorer la participation citoyenne par exemple). Elle est donc également bénéfique à l'économie nationale et permet même aux pouvoirs publics de réaliser des économies grâce à l'invention de nouveaux services à la population par des entreprises privées ou des citoyens engagés : par exemple, grâce à la réutilisation des données gouvernementales d'accidentologie, l'entreprise Groupama propose une plateforme pour choisir le trajet le plus sûr en voiture²⁴. Pour l'Etat, c'est un outil supplémentaire pour diminuer le nombre d'accidents sur les routes : la population est mieux protégée et moins de frais de soins doivent être engagés pour le système de sécurité sociale français. Les administrations croisant de plus en plus de fichiers pour faire des recoupements en matière de distribution des aides sociales, ou encore en matière de fiscalité, la data est également un moyen de contrôle pour le gouvernement. Mais au-delà de tout cela, l'open

²⁴ THEVENIN L. *Groupama exploite l'open data pour la sécurité routière* [en ligne]. Les Echos.fr (20/04/17) Disponible sur : < https://www.lesechos.fr/20/04/2017/LesEchos/22429-132-ECH_groupama-exploite-l-open-data-pour-la-securite-routiere.htm > (consulté le 23/06/18)

data est aussi un outil pour gagner la confiance des citoyens. L'ouverture participe de la transparence de l'action publique, or en informant mieux la population, celle-ci est sensée adhérer plus facilement à des projets collectifs. Des liens de confiance se créent et à terme des démarches de co-construction peuvent être envisagées. Ainsi, au fil du temps, l'ouverture des données va devenir un objet important aux yeux du législateur, qui va l'encourager et même la rendre obligatoire dans certaines collectivités.

L'Etat a besoin de la donnée car il est le premier à l'utiliser. Cette donnée sert à la réflexion, à la mise en place des politiques publiques mais aussi à leur évaluation. La collecte de ces données évite aux autorités de multiplier les enquêtes, et donc d'économiser, mais aussi d'avoir des données établies à l'échelle qu'elles souhaitent. D'autre part, certains domaines comme celui de l'aménagement urbain sont complexes à appréhender, de par la multiplicité des acteurs qui y prennent part, mais aussi les enjeux forts et changeants des territoires. Le partage des données, entre entreprises, collectivités, associations... doit alors être facilité au maximum afin de favoriser un travail transversal entre ces structures. Le renouvellement urbain d'un quartier prioritaire peut illustrer cette complexité. Par exemple, lors du réaménagement du quartier et du choix du pourcentage de logements sociaux, des partenaires tels que la CAF (Caisse d'Allocations Familiales) ou Pôle Emploi, peuvent être associés au projet en partageant leurs données. Cette coopération permettrait de réaliser des choix plus judicieux car plus adaptés au profil sociologique du quartier, dont ces structures publiques, présentes sur le territoire, ont une connaissance fine. Toutefois, pour garantir ce partage de données et leur utilisation il faut des « données de référence ». L'Etat doit donc garder la main sur certains référentiels afin que les utilisateurs, mais aussi lui-même, puissent s'y retrouver. Par exemple, l'Etat doit essayer de faire en sorte que le numéro SIREN des entreprises reste le moyen de les identifier, car si un acteur privé commence à changer ce référentiel en utilisant une autre information pour reconnaître les entreprises, alors il deviendra difficile d'aller trouver de nouvelles informations à ce sujet et de les croiser. La donnée doit aussi être de qualité et mise à jour régulièrement et être consultable constamment, et si possible en temps réel, via des API. Enfin, il faut garder à l'esprit que la donnée revêt un enjeu stratégique. L'exemple du GPS est à ce titre évocateur: développé dans un premier temps grâce à des satellites américains et destiné à l'armée, il représente

aujourd'hui un système indispensable pour bon nombre d'entreprises de l'aviation, des transports etc. et les autres pays, dont la France, tentent de déployer leurs propres satellites afin de ne plus être dépendants des Etats-Unis. Finalement, la valeur de la donnée réside dans « sa réutilisation, plus encore que dans son utilisation première »²⁵.

Les étapes clés de l'ouverture de la donnée :

Comme cela a été exposé plus haut, l'ouverture de la donnée constitue un nouvel enjeu pour les collectivités locales, puisqu'elle est devenue une obligation légale. Cette ouverture passe par plusieurs étapes²⁶, afin de bien préparer la donnée et de régler les détails qui accompagnent son ouverture :

1. « Identifier les données » :

Elles sont partout et nombreuses, c'est pourquoi le premier travail consiste à les identifier. C'est aussi le moment de regarder sous quel format elles peuvent être exportées, car elles sont parfois stockées dans des logiciels qui ne proposent pas d'exportation dans des formats plus ou moins universels. La sélection se fait souvent « par opportunité », les collectivités cherchant à « se distinguer à travers l'ouverture d'un jeu de données que personne n'a encore diffusé »²⁷

2. « Préparer les données » :

Les données doivent être brutes, or souvent elles sont modifiées, agrégées à d'autres données... car elles étaient utilisées par un service pour lequel il avait été jugé bon de procéder

²⁵ Secrétariat général pour la modernisation de l'action publique. *Rapport au premier ministre sur la gouvernance de la donnée 2015 – la donnée au service de la transformation de l'action publique* [en ligne]. (décembre 2015) Disponible sur : <

https://www.gouvernement.fr/sites/default/files/document/document/2016/01/rapport_au_premier_ministre_sur_la_gouvernance_de_la_donnee_2015.pdf > (consulté le 01/06/18)

²⁶ PRIOL J. *Le big data des territoires – Open data, protection des données, smart city, civic tech, services public... Les nouvelles stratégies de la donnée au service de l'intérêt général*. Buijsière : éditions FYP, 2017, 223 p.

²⁷ GOËTA S., MABI C. L'open data peut-il (encore) servir les citoyens ? *Mouvements* [en ligne]. 2014, n°79, pages 81 à 91. Disponible sur : < [https://www.cairn.info/revue-mouvements-2014\(3\)page-81.htm](https://www.cairn.info/revue-mouvements-2014(3)page-81.htm) > (consulté le 20/06/18)

ainsi. Certains jeux comportent des erreurs, or les données ouvertes doivent être exactes et à jour. La question de la protection des données personnelles intervient à cette étape : il faut anonymiser les bases à ouvrir, mais garder à l'esprit que cela peut ne pas suffire, car à force de croiser et recroiser des données, il est possible de ré-identifier une personne.

3. « Publier les données » :

La première question à poser est de savoir si la collectivité souhaite être la seule à contribuer à l'alimentation de sa plateforme open data ou si elle va collaborer avec d'autres acteurs. Les municipalités qui choisissent de travailler seules sont plus visibles, mais cela représente un coût, alors que d'autres plateformes peuvent être mises en place au niveau départemental par exemple et ainsi mutualiser des moyens financiers et humains. Le choix du format des données entre aussi en compte : un fichier « .csv » correspond à un fichier « .xls », mais ce dernier ne peut s'ouvrir que sur le logiciel Excel, qui est payant. Enfin, la dernière décision à prendre est de choisir entre un partage de la donnée au fur et à mesure des mises à jour effectuées par la collectivité, ou le partage en temps réel, qui demande l'utilisation d'interfaces de programmation applicatives (API).

4. « Choisir une licence » :

La question qui se pose est : « peut-elle être utilisée par tout le monde, gratuitement et pour n'importe quel usage ? » (J. Priol, 2017) et donc est-ce que des entreprises privées peuvent utiliser cette donnée sans contrepartie ? La licence ODbL est contraignante pour les réutilisateurs privés à but lucratif, puisqu'elle oblige à repartager les travaux réalisés à partir des données récupérées. La licence ouverte permet d'utiliser des données même à des fins commerciales, elle oblige seulement à mentionner l'origine des données utilisées.

5. « Promouvoir sa démarche » :

Il convient de communiquer en amont et en aval sur l'ouverture des données, afin de susciter un intérêt dans l'objectif de faire en sorte que les données libérées soient réutilisées. Il existe 4 types de réutilisation :

- La simple consultation par la population ;

- La médiation : pour comprendre des politiques publiques. Cela passe par des outils de visualisation ou encore des cartes interactives ;
- Des données destinées à des applications et donc mises à jour en temps réel : elles nécessitent de passer par une API ;
- Des données plus spécialisées, destinées aux entreprises ou à des personnes initiées qui souhaitent par exemple développer des applications.

Afin de poursuivre la politique open data dans le temps, il faut également animer la démarche, afin de susciter de l'intérêt pour les données ouvertes. Cela peut se faire à travers des concours, des hackathons (« marathons de programmation » (J. Priol, 2017)) lancés par la collectivité. Des concours sont organisés sur quelques jours : les participants peuvent venir de la société civile, de la sphère associative ou privée selon les règles fixées. L'objectif est de produire une application, un outil ou de nouvelles données en un temps restreint, avec à la clé une remise de prix. Les applications créées permettront de rendre service à la ville en améliorant la cadre de vie des citoyens. A Toulouse, la métropole s'est associée à Intel, Orange et la French Tech Toulouse (communauté de startups toulousaines) pour créer un hackathon, portant sur le thème de la ville intelligente et de la « smart citoyenneté »²⁸. Cette manifestation, qui intégrait uniquement des élèves ingénieurs en informatique (provenant d'Epitech Toulouse ou encore de l'INSA), a été organisée dans le but de développer des services à partir d'objets connectés. La première place a été remportée par le projet « Responcivility » : grâce à des capteurs, cette application permet de détecter les stationnements illicites sur les places réservées aux personnes à mobilité réduite et sur les passages piétons. Le projet « EcoStreet » permet de mieux organiser les tournées de collecte des déchets ménagers en se basant sur le remplissage des containers. Enfin, le projet « SAINT » détecte les tentatives d'agression, grâce à des capteurs de son dans les rues qui remontent les informations collectées à une plateforme d'assistance.

²⁸ Toulouse Métropole. #CP – Des élèves d'EPITECH Toulouse remportent le hackathon smart city de Toulouse [en ligne]. (18/12/15) Disponible sur : < <https://www.toulouse-metropole.fr/-/cp-test-des-eleves-de-epitech-toulouse-remportent-le-hackathon-smart-city-de-toulouse> > (consulté le 28/05/18)

En 2017, 171 collectivités territoriales ont un portail de données ouvertes ou publient leurs données sur le site open data du gouvernement (data.gouv.fr)²⁹. Parmi ces collectivités territoriales on compte 11 régions et 32 départements. Il y a finalement, presque 1 700 communes qui se sont lancées dans l'open data en France aujourd'hui.

²⁹ PRIOL J. *Le big data des territoires – Open data, protection des données, smart city, civic tech, services public... Les nouvelles stratégies de la donnée au service de l'intérêt général*. Buijsière : éditions FYP, 2017, 223 p.

2. L'UTILISATION DE LA DONNÉE A TOULOUSE METROPOLE

2.1. Toulouse, parmi les pionniers dans l'open data

La donnée à Toulouse : une gestion de plus en plus complexe pour répondre aux attentes du législateur

L'informatisation de l'administration a débuté dans les années 1970 à Toulouse, il y a donc plus de 40 ans. Aujourd'hui quasiment l'ensemble des services de la métropole est informatisé, mis à part certains services qui agissent directement sur le terrain (type espaces verts). Il y a encore quelques années la donnée n'était pas envisagée en tant qu'objet. Les services prenaient seulement en compte l'outil, soit le logiciel. Or depuis une quinzaine d'années, quand on change d'outil on se pose la question de la reprise des données. La collectivité commence donc aujourd'hui à cartographier la donnée.

Cartographier la donnée : pour quoi faire ?

La cartographie des données permet de visualiser toutes les données dont on dispose et leur provenance. Ainsi, toute la collectivité peut avoir une vision globale de ce dont elle dispose.

Afin d'être en conformité avec la nouvelle RGPD, il est aujourd'hui obligatoire de réaliser une cartographie des données, notamment celles relatives à la vie privée. Cette opération consiste à recenser de manière précise l'ensemble des données dont on dispose, leurs traitements, leurs flux, les composantes par lesquelles elles circulent et les acteurs qui les traitent³⁰. Cela oblige également les détenteurs de ces données à se préoccuper d'autant plus de la sécurisation de ces données.

³⁰ Datagalaxy.com. *Cartographie des données personnelles : l'indispensable chantier pour démarrer avec GDPR* [en ligne]. (2018) Disponible sur : < <https://www.datagalaxy.com/blog/gdpr/cartographie-donnees-personnelles-commencer-gdpr/> > (consulté le 14/06/18)

Cela est très compliqué à réaliser, comme l'explique Sandrine Mathon, chef du service administration la direction des systèmes d'information de la ville de Toulouse et de Toulouse Métropole, responsable de l'Open Data de la Métropole³¹ : « Là c'est un autre boulot... Déjà la cartographie des outils, sachant que l'on a à peu près 430 applications (...) c'est déjà assez conséquent. (...) donc maintenant il faut que dans chaque outil on arrive à connaître les types de données qui sont enregistrées. Et ça c'est un énorme travail (...) à détailler appli par appli le genre de donnée qu'il y a dedans. ». Chaque donnée doit donc être repérée et des questions d'accès à ces données se posent. Par exemple, sur le logiciel qui gère les traitements des fonctionnaires, certaines informations sont plus sensibles que d'autres : données de santé, mandats syndicaux... Dans la cartographie des données du logiciel de payes, la collectivité doit donc déterminer qui a accès à quelles informations. Pour ces données, considérées comme sensibles, le niveau d'accès sera donc restreint à certains membres du personnel et cela doit être précisé dans la cartographie des données.

Toulouse Métropole, dans les pionniers de l'Open Data en France :

C'est dans les années 2010 que le mouvement de l'open data est lancé en France. Un climat propice à ce phénomène va se mettre en place, grâce à des volontés politiques qui émergent dans de grandes villes comme Nantes, Rennes, Brest ou encore Toulouse, mais aussi à des associations très actives comme LiberTIC. La multiplication des rencontres entre ces différents acteurs et leur constitution en un réseau national, avec l'association OpenDataFrance, va permettre de lancer réellement le mouvement open data en quelques années.

Au début de l'année 2011, une volonté politique émerge de lancer l'ouverture de la donnée publique à Toulouse Métropole. Les objectifs initiaux qui sont exposés aux élus sont les suivants :

- La transparence administrative ;

³¹ Entretien réalisé en juin 2018 avec Sandrine MATHON, chef du service administration à la direction des systèmes d'information de la ville de Toulouse et de Toulouse Métropole, responsable de l'Open Data de la Métropole.

- Le développement économique ;
- Le marketing territorial.

Entre janvier et avril 2011, un temps de veille et de collecte est donc organisé. Tout commence grâce à une personne qui travaillait à la DSI, qui connaissait le chef de projet qui traitait l'open data à la ville de Paris. La DSI s'est donc rapprochée de cette personne, mais également d'une association Nantaise : LiberTIC, qui aidait la ville de Nantes, mais aussi le département et la région Loire-Atlantique dans le lancement de l'open data. La ville de Toulouse se met donc en relation avec ces personnes. La personne contactée dans l'association Libertic étant très active sur le sujet, cela a donné accès à la ville de Toulouse à tout un réseau pour se lancer dans l'ouverture de la donnée publique. En mars 2011 les élus de Toulouse Métropole décident officiellement de se lancer dans l'open data.

La FING (Fédération Internet Nouvelle Génération) organise une conférence à Marseille en juillet 2011. C'est une nouvelle occasion de créer de nouveaux contacts. Une association toulousaine, nommée Ecosysteme, va aussi jouer un rôle déterminant à travers une lettre ouverte qu'elle adresse au maire de Toulouse de l'époque, Pierre Cohen, demandant de libérer les données.

Un terrain fertile naissait donc en faveur de l'ouverture des données publiques à l'échelle nationale à cette époque. Cette effervescence autour de l'open data, aussi bien sur l'ensemble du territoire national, qu'avec l'appel du collectif citoyen toulousain puis le soutien politique des élus, va permettre d'obtenir l'accord du Directeur Général des Services de la métropole de Toulouse et donc confirmer le lancement de l'open data à Toulouse.

Toutefois, le phénomène open data n'en étant qu'à ses balbutiements, la métropole avance à tâtons, mais veut faire vite. Pour agir de manière efficace, tout commence par un travail avec le service de la communication, qui aide la DSI à construire une plateforme pour publier les données. La première version du site était donc basique, avec un simple bouton « télécharger » pour accéder aux données. Parmi les premières données publiées on retrouve : l'altimétrie de voies, les pôles territoriaux, les mairies de quartier ou encore les quartiers de la démocratie locale.

Des échanges s'opèrent entre les différents chefs de projet open data, une dizaine à l'époque, ce qui a facilité la communication entre eux.

Dans un second temps la question des licences de réutilisation s'est posée. Un travail s'engage alors avec OpenStreetMap. Cette association était intéressée par la récupération de données géographiques (qui appartenaient à la métropole) qui concernaient les filaires de voiries, afin d'ouvrir la carte OSM (OpenStreetMap) de Toulouse. Après avoir passé l'été à travailler avec eux et à discuter des licences de réutilisation, le choix s'est porté sur la licence ODbL pour l'ouverture des données. Cette licence a été choisie car c'était celle qui était la plus ouverte à cette époque, et que la licence ouverte d'Etalab n'existait pas. De plus, ODbL oblige au repartage à l'identique, ce qui selon Sandrine Mathon « permet que tout le monde participe au commun » et représente donc un grand avantage.

En 2013 Sandrine Mathon ainsi que d'autres personnes qui travaillent sur l'open data dans les collectivités rencontrent Henri Verdier, un entrepreneur toulousain spécialisé dans le numérique qui vient d'être nommé directeur d'Etalab (et sera nommé Administrateur Général des Données par le gouvernement un an plus tard). Sur conseil d'Henri Verdier, afin d'asseoir sa légitimité démocratique et d'être entendu par les instances nationales, le mouvement open data décide de monter une association. Finalement, c'est à Toulouse qu'elle sera créée et où son siège y sera implanté en 2013. Cette association, présidée par l'élue adjoint au Maire de Toulouse Bertrand Serp jusqu'en novembre 2017, rassemble les collectivités qui se lancent dans l'open data et accompagne celles qui souhaitent rejoindre le mouvement, mais aussi des entreprises qui veulent entrer dans cette logique.

Toulouse se situe aujourd'hui parmi les collectivités les plus en pointe dans le domaine de l'open data, de par son statut de « ville pionnière ». Elle oriente l'open data vers le développement économique. L'open data sert aussi d'outil pour la participation citoyenne, c'est donc la DAT (Direction de l'Action Territoriale), qui est en charge de la participation citoyenne, qui va prendre l'initiative de s'appuyer sur l'open data pour travailler son action. Certaines collectivités, comme Rennes, sont très axées sur la médiation citoyenne. La ville décide d'un budget participatif grâce à l'open data, elle organise également des ateliers de la donnée auprès des écoles... Elle est très active sur le pan participation citoyenne. Lyon est plus axée sur le développement économique.

Les finalités de la mise à disposition de la donnée peuvent donc différer selon les contextes territoriaux, les ambitions et les projets des élus, car cette démarche est en lien avec une volonté politique. L'open data, émanant d'une vision politique, peut donc orienter l'action publique sur un territoire : développement économique, participation citoyenne... Chaque collectivité choisit son champ d'action, qu'elle développera en s'appuyant sur sa politique open data.

La donnée ouverte va donc influencer le travail et les choix de l'aménageur, qui doit prendre en compte de nouvelles informations amenées par la sphère citoyenne ou d'autres acteurs, qui entrent en jeu à partir du moment où les données sont ouvertes. Il est donc obligé de composer avec ce nouvel outil, qui va enrichir les projets par la finesse des informations que cela va apporter, mais aussi les complexifier, notamment par les avis divergents d'acteurs multiples. Il devra alors trancher, dans l'intérêt général, pour mener à bien le projet qu'il porte. Tout en tenant compte de ces nouvelles données, il devra faire attention de ne pas s'y noyer.

Transformation de l'administration et développement de nouveaux métiers

Les Etats-Unis se démarquent par leur gouvernance de la donnée, qui définit aujourd'hui de nouveaux services municipaux et de nouvelles fonctions en lien avec les données. La fonction de *chief data officer* (CDO) a été initiée sur ce continent, elle est aujourd'hui appliquée en France, sous le nom d'administrateur général des données. Le *chief data officer* français est Henri Verdier, nommé en 2014 et aujourd'hui à la tête d'Etalab. Tout comme outre Atlantique, les CDO commencent à se déployer également en région en France. Le *chief data officer* est celui qui orchestre les actions en lien avec l'open data dans la ville. Il a également en charge l'évaluation des politiques publiques grâce aux données récoltées, ainsi que le croisement des données afin d'anticiper les phénomènes urbains. Son service est en lien avec le service dédié au numérique. Il travaille en transversalité avec les autres services municipaux, car la démarche même de l'open data a pour objectif de décroiser les services pour favoriser l'échange et le croisement des données.

L'open data nécessite donc une réorganisation des administrations, mais aussi de nouveaux métiers, comme celui d'éditeur de données, qui aide à la production de la data.

2.2. Gestion et échanges de données : entre liberté de l'ouverture des données et respect des licences

La gestion de l'Open Data à Toulouse :

A Toulouse Métropole, c'est la Direction des Systèmes d'Information (DSI) qui s'occupe de l'ouverture des données. Elle les publie sur la plateforme « data.toulouse-metropole.fr ». Lorsque la métropole demande à ce qu'une donnée soit publiée (par une volonté politique ou d'un service) la DSI se rapproche du service concerné par cette donnée et communique avec afin de la récupérer. En même temps, cela permet à la DSI de collecter d'autres jeux de données qu'il peut être utile d'ouvrir. Pour Sandrine Mathon, l'ouverture d'une donnée se fait au fur et à mesure : « C'est très pragmatique. Il n'y a pas de plan. C'est : qu'est-ce qu'on a ? Qu'est-ce qu'on nous demande ? Et comment les services s'y prennent ? ». Le choix et l'ouverture de la donnée se fait au jour le jour, selon les opportunités.

Finalement, l'ouverture de la donnée se fait au coup par coup. Cela s'explique d'une part par le fait que la démarche vienne de la collectivité : le législateur n'intervient pas, il n'y a donc pas de procédure à suivre qui soit décrite par l'Etat. La collectivité est seule responsable de l'ouverture de ses données, c'est à elle de décider ce qu'elle publie et comment elle procède pour y parvenir. L'ouverture se fait petit à petit, selon la disponibilité des services, les décisions politiques ou celles prises par le service de la DSI pour animer le projet open data, ou encore selon les envies particulières qu'ils peuvent avoir de partager certaines données. Tout le monde tente d'avancer en s'adaptant au rythme de chacun car, surtout au commencement, il fallait faire comprendre aux agents l'intérêt de cette démarche, puis les former à la manière d'ouvrir un jeu de données.

Illustration 4 : Données publiées sur la plateforme open data de Toulouse Métropole (source : data.toulouse-metropole.fr)

La validation des données, passe par les services de la DSI, mais les élus peuvent être sollicités pour confirmer l'ouverture de certaines données si le service un a doute. Toutefois, avec la loi pour une République Numérique, la collectivité se questionne de moins en moins puisqu'elle doit publier le maximum d'informations. Les seules données à censurer ou à bien retravailler avant de les publier sont celles qui peuvent permettre une ré-identification. Cela a été le cas avec un jeu de données sur les abonnés à la médiathèque, qui renseignait le nombre d'abonnés par rue de Toulouse. La décision a été prise de censurer ces données pour des questions de maintien de l'anonymat.

Sandrine Mathon note également que certains services se sont très bien adaptés au mouvement open data : « par exemple les archives se sont très bien approprié ce sujet. Ils ont leur propre compte pour publier tout seuls leurs données sur le site Open Data Toulouse Métropole. Les musées aussi ils sont presque prêts à publier tous seuls (...) les finances aussi publient tous seuls... ». La DSI apporte une aide aux services qui souhaitent publier des données pour leur première publication, afin de leur montrer comment fonctionne la plateforme, et qu'ils puissent être autonomes par la suite. D'autre part, elle a remarqué que

certaines publications répondent à des besoins internes : par exemple l'accessibilité des bâtiments aux personnes à mobilité réduite a été cartographiée directement en open data.

Enfin, la publication des données peut être parfois complexe. Quand le service concerné possède déjà son jeu de données sur un tableur tout est plus simple, il n'y a qu'à l'insérer sur la plateforme. Par contre, s'il faut consolider le tableau en agrégeant des données d'autres services cela prend alors beaucoup plus de temps. Par exemple, sur l'accessibilité des bâtiments aux PMR, un agent a passé une dizaine de jours à agencer les données pour les rendre publiables. La structure du jeu de données à publier est donc une étape clé de la démarche mais une étape très chronophage. Cela est d'ailleurs l'une des causes majeures de la faible ouverture de données massives de la part des collectivités³².

Echanges de données et coopérations :

Ce sont surtout les acteurs privés qui sont intéressés pour récupérer les données de la ville, comme Waze par qui souhaite obtenir des informations sur les fermetures de routes, les travaux à venir... tout cela dans le but d'enrichir son application. Mais Sandrine Mathon précise que la licence ODbL utilisée pour les données publiques à Toulouse doit être scrupuleusement respectée. Toutefois des dérogations peuvent être faites : ainsi, l'entreprise qui utilise un jeu de données en ODbL n'est pas obligée de repartager la donnée en licence ODbL et conserve donc ces données fermées. C'est le cas avec l'entreprise TomTom qui, après délibération du Conseil Métropolitain en juin, peut maintenant utiliser les données ouvertes en dérogation de la licence ODbL car elle apporte un service particulier à la ville : La société injecte, à la demande de la métropole, des règles pour le calcul des itinéraires de leurs utilisateurs. Pour les autres entreprises, qui ne sont pas prêtes à fournir un service comme celui-là, aucune dérogation à la licence ne sera délivrée (c'est le cas pour Waze à ce jour). Pour Google, la question a été résolue au niveau national. Google souhaitait accéder aux données Tisséo et à celles d'autres opérateurs de transports collectifs urbains dans différentes grandes

³² GOËTA S. Un air de famille : les trajectoires parallèles de l'open data et du big data. *Informations sociales* [en ligne]. 2015, n°191, pages 26 à 34. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-26.htm> > (consulté le 20/06/18)

villes de France. La mission Etalab est donc en train de créer une plateforme dédiée à ces données (transport.data.gouv), sur laquelle toutes les données liées au transport seront centralisées.

Illustration 5 : Jeux de données disponibles sur le site transport.data.gouv (source : transport.data.gouv)

Jeux de données valides disponibles



GTFS des Transports Urbains du Grand Avignon – TCRA

Liste des arrêts, horaires et parcours de tous les bus du réseau TCRA circulant sur le territoire de l'agglomération du Grand Avignon.

[Voir sur data.gouv.fr](#)
[Télécharger](#)

Territoire **Grand Avignon**

Format **GTFS**

État **Valide**



GTFS - Strasbourg

Fichiers GTFS du réseau de transport de l'Eurométropole de Strasbourg. Ne couvre pas les transports du Bas Rhin (CTBR)

[Voir sur data.gouv.fr](#)
[Télécharger](#)

Territoire **Eurométropole de Strasbourg**

Format **GTFS**

État **Valide**



Données sur le réseau de bus SURF (Service Urbain de la Région Fougèraise)

Trois fichiers sont inclus dans ce jeu de données :

[Voir sur data.gouv.fr](#)
[Télécharger](#)

Territoire **Fougères Agglomération**

Format **GTFS**

État **Valide**

Toutefois la licence ODbL devra être respectée. D'autres entreprises, comme Waze, sont en discussion avec la ville, mais ils ne souhaitent pas aujourd'hui modifier leurs itinéraires selon les demandes de la ville et sont donc obligés de respecter la licence ouverte utilisée par la ville.

Afin d'avancer dans la libération des données, Toulouse Métropole continue d'être en réseau avec les adhérents d'OpenDataFrance mais aussi avec le département, la région, le SICOVAL et la préfecture de région. Toutes ces structures, qui libèrent de la donnée, permettent de s'enrichir les unes les autres.

Plus largement, OpenData Lab incite les structures publiques à ouvrir leurs données pour qu'elles soient réutilisées. La carte qui suit est un exemple d'utilisation des données ouvertes de la DRAAF (Direction Régionale de l'Alimentation de l'Agriculture et de la Forêt) Occitanie, mises à disposition depuis la plateforme OpenData Lab. Elle représente les Zones de Revitalisation Rurale (ZRR) en 2014, dans la région Occitanie. Les ZRR sont des aides fiscales pour favoriser l'installation d'entreprises sur des territoires identifiés comme fragiles. Le

périmètre des ZRR ayant été redéfini en 2017 (les données ne sont pas encore disponibles), cette carte peut permettre d'alimenter le débat sur le redécoupage de ces zones qui a fait grand bruit auprès de la population à sa sortie.

L'ouverture de la donnée se fait donc selon les volontés de la métropole, le législateur n'encadrant pas la démarche, la collectivité avance pas à pas. Elle doit également gérer les coopérations créées par l'open data : certaines seront stratégiques, notamment celles avec des acteurs privés (comme Waze), d'autres permettront d'obtenir de l'aide dans cette démarche, comme cela peut être le cas auprès de l'Open Data Lab de la région Occitanie.

2.3. Retour d'expérience sur le stage de master 1 à l'Observatoire de la Métropole : faire face aux obstacles en amont de l'ouverture de la donnée

Aujourd'hui tournée vers l'ouverture de la donnée, la métropole toulousaine n'a cependant pas encore réussi à résoudre tous les désagréments qui peuvent y être liés. C'est ce que j'ai pu expérimenter au cours de mon stage à l'Observatoire de Toulouse Métropole, un service qui travaille la donnée, ouverte ou non, au quotidien.

Ce service est rattaché à la délégation à l'aménagement et travaille trois publications annuelles sur Toulouse et la métropole :

- Le « Toulouse Métropole en Chiffres » : à l'échelle de Toulouse, il dresse le portrait de la ville et livre des informations détaillées sur de multiples thèmes (économie, logement, environnement...) ;
- Le « Mémento Métropolitain » : livre le portrait des 37 communes que compte la métropole toulousaine à travers des données statistiques ;
- Le « Mémento Toulouse les quartiers de la démocratie locale » : présente les quartiers toulousains à travers des données statistiques et rassemble plusieurs cartes sur le logement à Toulouse.

Ces publications sont tirées au format papier et distribuées aux services de la métropole, leur apportant des connaissances chiffrées et cartographiées. Les publications du service de l'Observatoire sont utiles à l'ensemble des agents. Après chaque publication, de nombreux services réclament ces documents pour travailler, car ils regroupent des informations générales mais aussi des chiffres précis, qui alimentent le travail et les réflexions. Les publications sont également disponibles en ligne sur le site de la métropole, à destination du grand public. Ces documents apportent des informations à toute personne, citoyen, étudiant, universitaires etc. désireuse d'en connaître d'avantage sur Toulouse et sa métropole.

Au cours du stage, j'ai eu en charge deux missions principales. La première a été de mettre à jour le réseau des référents de mon service. En effet, pour constituer ses publications annuelles, l'Observatoire entretient des relations avec l'ensemble des services de la métropole. Il est donc primordial pour ce dernier de garder une liste de contacts à jour afin de récolter chaque année de nouvelles données actualisées. La seconde mission portait sur la mise à jour des publications du service, ainsi qu'une étude sur les permis de construire et déclarations préalables déposés dans la ville de Toulouse. J'ai donc eu l'occasion de traiter différentes bases de données et d'approcher les difficultés auxquelles le service peut être confronté de sa récolte, à son traitement et son analyse.

Pour élaborer ses publications, l'Observatoire remet chaque année ses données à jour. Toutefois les agents de ce service peuvent se heurter à des difficultés liées à la donnée lors de leur travail : c'est ce qui sera exposé dans cette partie. Les obstacles auxquels doit faire face la métropole pour ouvrir ses données sont de différentes natures : techniques, politiques... Enfin, il convient de garder à l'esprit que pour être ouvert, un jeu de données doit être vérifié et réutilisable par tous.

La première difficulté à laquelle j'ai été confrontée a été celle d'accéder au serveur où se trouvaient des bases de données que je devais exploiter, afin de mettre à jour une carte du Mémento Métropolitain sur les Déclaration Attestant Achèvement et Conformité des Travaux (DAACT). La métropole possède plusieurs serveurs, or mon service se trouve dans le serveur X alors que les données dont j'avais besoin se trouvaient dans le serveur Y qui héberge le service des autorisations d'urbanisme. Avec l'aide de mon chef de service, nous avons donc

demandé un accès à ce serveur. D'autres données ont directement été récupérées auprès de personnes connues par mon chef de service. Finalement, la première difficulté lorsque l'on souhaite récupérer une donnée est de savoir où la trouver et qui contacter pour la récupérer.

Avant de traiter les données, il faut déjà se poser la question de la compatibilité des fichiers que l'on va utiliser : Excel, libre office... pour savoir si l'on va pouvoir les ouvrir sur notre ordinateur, car certains PC ne possèdent pas des logiciels payants comme Excel. De même, lorsque j'ai travaillé sous le logiciel de cartographie Mapinfo, les fichiers Excel peuvent être entrés dans le logiciel, mais les tables de données produites par le logiciel ne peuvent pas être extraites au format Excel. Il faut les enregistrer au format Dbase afin de pouvoir les ouvrir, uniquement avec libre office. En amont se pose donc la question des formats des jeux de données et de leur compatibilité avec les différents logiciels de travail.

Un autre obstacle auquel j'ai été confrontée est celui de la mise en forme des données. Plusieurs questions se posent lorsque des données sont publiées, même lorsqu'elles sont déjà traitées sous forme de cartes ou autre, comme c'est le cas dans les publications du service de l'Observatoire. A qui s'adressent ces données ? Qui va les consulter ? Dans quel but ? Car un autre problème, auquel peuvent être confrontées les personnes qui s'intéressent à une base de données, est celui de leur appropriation. Celle-ci peut parfois poser problème entre des services du même domaine, qui n'ont pas les mêmes manières de travailler.

Lors de l'utilisation du logiciel SIG (ici, Mapinfo), j'ai géocodé³³ la table de données des DAACT. Pour y parvenir, j'ai utilisé la table nommée « voie_numerue » qui répertorie l'ensemble des adresses de la ville. En géocodant la table des DACCT avec l'aide de « voie_numerue », certaines lignes n'ont pas été géocodées : leur adresse n'a pas été retrouvée. Cela est dû la plupart du temps à des erreurs de mise en forme de l'adresse dans les tables utilisées. En effet, pour croiser une adresse avec la table « voie_numerue » il faut qu'elle soit conforme au format suivant : 0012 (numéro de voie) RUE (type de voie) DES LILAS

³³ Géocodage : consiste à affecter des coordonnées géographiques (longitude/latitude) à une adresse postale (source : <https://fr.wikipedia.org/wiki/Géocodage>)

(nom de la voie). De plus, il existe plusieurs possibilités pour les types de voie : AV (avenue), ALL (allée), RUE, CHE (chemin), PAS (passage)... Lorsque l'on récupère un jeu de données, issu des services internes à la métropole les adresses sont en général sous ce format, mais lorsque les jeux utilisés ne proviennent pas de la métropole, les adresses peuvent prendre une autre forme. Il faut donc toutes les remettre au bon format, or cette opération peut prendre énormément de temps. Enfin, avec les outils informatiques, la moindre erreur d'orthographe, de virgule ou de numéro met en péril le géocodage des lignes de la table.

La mise en forme peut donc poser problème sur la donnée en elle-même, mais aussi sur la construction du tableur. Par exemple, lorsque j'ai voulu croiser des données provenant de bases PERVAL sur plusieurs années (base des transactions immobilières), cela s'est avéré très compliqué. Je possédais cinq bases différentes, chacune correspondant à une année. Chaque jeu de données était construit différemment et pouvait contenir des informations en plus ou en moins. Je souhaitais garder un maximum d'informations car les jeux comprennent énormément de colonnes, avec des informations très riches, mais pour regrouper tous ces jeux sans perdre d'informations il fallait que je mette en forme chaque jeu pour qu'ils soient constitués de la même façon. Cela était donc extrêmement compliqué et je ne l'ai pas fait. Or cela a compliqué toutes les tâches suivantes effectuées avec ces données, que j'ai dû croiser avec d'autres données pour chaque année PERVAL, ce qui m'a demandé au final de gérer un grand nombre de fichiers.

Enfin, un dernier problème est celui de la vérification des données. En utilisant une base qui concerne les Déclarations Préalables (DP), j'ai voulu calculer la moyenne des surfaces plancher des projets. La moyenne aurait dû être inférieure à 40 m², soit la surface maximale autorisée pour une DP, mais elle dépassait largement ce chiffre. Ce résultat était donc forcément faux. En regardant de plus près les chiffres des surfaces des projets, je me suis rendue compte que certains étaient faux. J'ai donc dû vérifier tous ces chiffres en retrouvant, pour chaque ligne (soit une centaine au total) les documents de la DP qui étaient scannés et stockés sur le serveur de la collectivité. Finalement, cette vérification a été chronophage et j'aurais pu sortir des résultats erronés si je n'avais pas prêté attention à ces chiffres, car je pensais qu'ils étaient justes.

Pour certaines données jugées sensibles, le choix est fait de ne pas les publier. Par exemple, dans le Toulouse Métropole en Chiffres, les données concernant les ZAC (Zones d'Aménagement Concertées) sont limitées car les élus sont parfois réticents à rendre certains chiffres publics. Lors du premier Atelier d'Observation Métropolitaine, auquel j'ai eu l'occasion de prendre part, une personne du service de la police municipale a rapporté que son service dispose d'un jeu de données sur la criminalité. Or, la métropole n'a aucun intérêt à publier ce genre de données qui pourraient donner une mauvaise image de la ville ou encore stigmatiser certains quartiers. De la même manière les données qui permettent la ré-identification d'individus ne sont pas publiées afin de préserver l'anonymat.

La donnée est souvent représentée comme une série de chiffres incompréhensible, ce qui décourage beaucoup de personnes de se les approprier lorsqu'elles sont publiées dans un format brut (type tableur). Pourtant, elle revêt une importance capitale dans la maîtrise de nos territoires aujourd'hui, à l'ère du numérique, dans une société hyper connectée qui produit de la donnée partout et en continu. Finalement, j'ai pu pointer plusieurs obstacles à l'ouverture des données durant mon stage :

- La récupération des jeux de données, qui peut prendre du temps et demande un réseau de contacts à avoir pour savoir auprès de qui les récupérer ;
- Le format des jeux de données ;
- La structuration des données, au niveau de l'information elle-même mais aussi dans la construction des tables ;
- La véracité des données ;
- Les questions de censure qui se posent en amont dans la publication de certaines données.

Au cours de l'entretien avec Sandrine Mathon, j'ai relevé un autre frein, apparu au lancement de l'open data dans la collectivité : certains services n'ont pas tout de suite compris quelle utilité pouvait présenter l'ouverture de la donnée. Finalement, ce préjugé s'est rapidement envolé. L'un des meilleurs exemples est celui de l'herboriste du muséum de Toulouse, qui a publié l'herbier des plantes sauvages, c'est-à-dire les « mauvaises herbes » qui

poussent en ville. Alors que plusieurs personnes peuvent ne pas en voir l'utilité, d'autres ont trouvé un intérêt à réutiliser ces informations et une application a été créée.

D'autres freins sont pointés par la communauté scientifique dans la diffusion de la donnée. Le fait que ce sont les agents des collectivités qui sélectionnent eux-mêmes les jeux à publier conduit à une restriction des données ouvertes, car ils doivent d'abord vérifier la donnée et peuvent vouloir éviter de se mettre en danger vis-à-vis de leur hiérarchie.

2.4. L'open data : de nombreux bénéfices et des résultats concrets pour améliorer la vie en ville

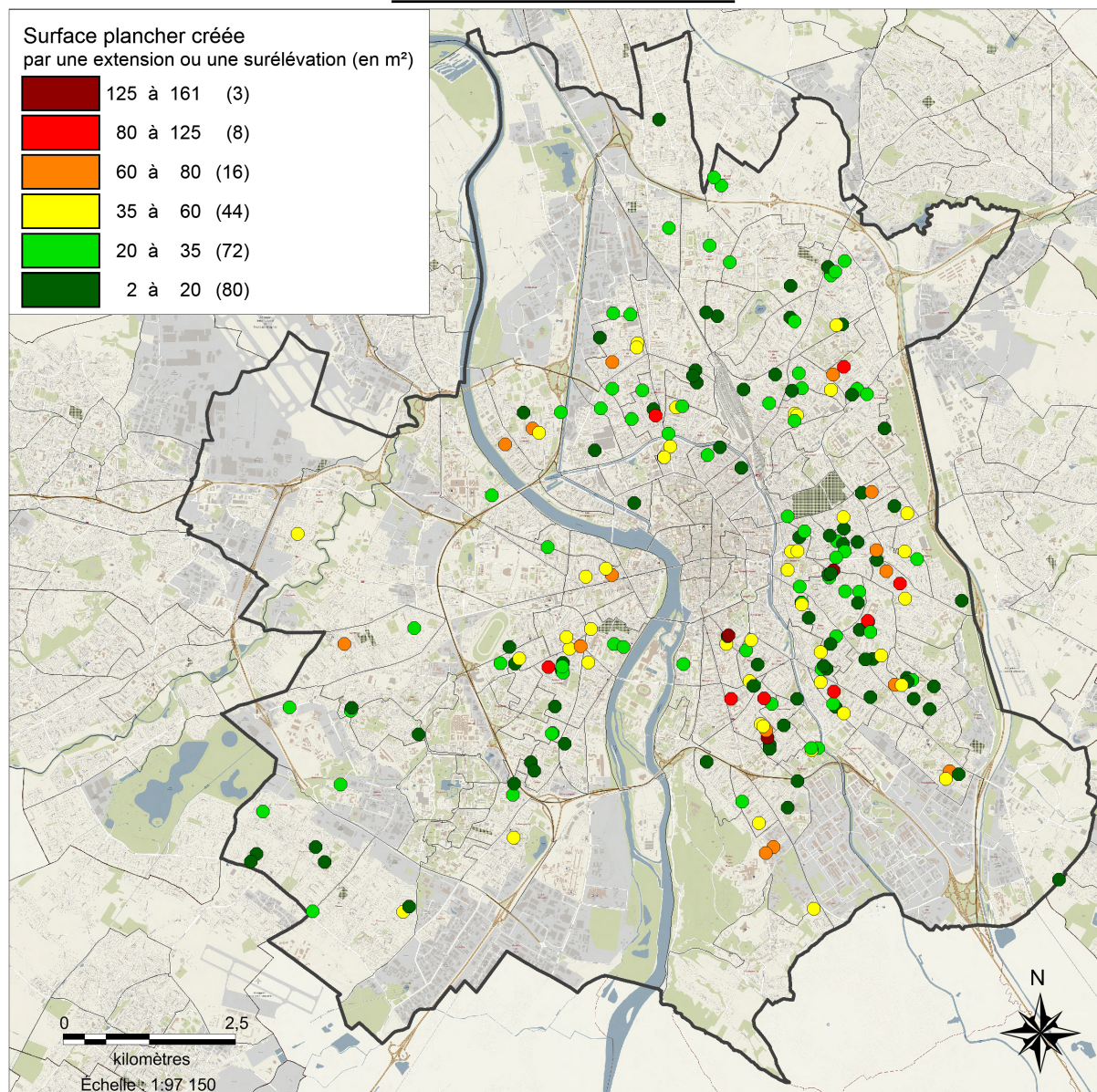
Les bénéfices liés à l'ouverture de la donnée à Toulouse Métropole :

Au cours du stage, j'ai pu constater que l'ouverture des données entre les services de la métropole peut permettre de créer de nouvelles connaissances et de nouveaux outils.

J'ai eu l'occasion de travailler sur le suivi des extensions et surélévations liées à des déclarations préalables ou à des permis de construire, dans la ville de Toulouse. La carte qui suit représente les surfaces créées par ces extensions et surélévations, entre 2013 et 2017. L'information que j'ai produite pourrait par la suite faire l'objet d'une appropriation par la métropole, afin de réfléchir à plusieurs objets comme la décision de réaliser un suivi de ce genre de travaux, qui est une manière d'intensifier la ville, mais également la possibilité de créer un nouvel impôt sur ces constructions qui constituerait une source de revenus supplémentaire pour la ville. Ce travail peut donc aujourd'hui avoir une réelle valeur pour la métropole si elle décide de s'en emparer, mais cela n'aurait pas été possible sans les échanges de données entre services, ce qui démontre l'importance et le grand avantage de l'open data.

Illustration 7

Surfaces plancher créées entre 2013 et 2017, dans des biens vendus entre 2010 et 2016 à Toulouse



Réalisation : Laura MANEVILLE, 2018 - Source : Toulouse Métropole et Ville de Toulouse, PERVAL

Selon Sandrine Mathon, l'ouverture de la donnée a eu plusieurs conséquences bénéfiques en interne : plus de transversalité au sein de la collectivité, mais aussi de la modernisation et de l'efficacité car les agents n'ont plus besoin de chercher les données, elles sont centralisées sur la plateforme et prêtes à l'emploi.

Un autre atout lié à l'open data, qui n'était pas forcément attendu au départ, est celui de l'amélioration de la qualité des données produites par les services. En effet, auparavant les

services produisaient une donnée de moindre qualité, mais ils se débrouillaient, conscients des marges d'erreurs de leurs jeux de données. A Lille, la collectivité a ouvert un jeu de données sur l'âge des défunts. Mais ils se sont rendus compte qu'un grand nombre de personnes mourraient soit entre 1 et 2 ans, soit entre 115 et 117 ans : il y avait donc une anomalie, puisqu'il ne pouvait pas y avoir une mortalité infantile aussi élevée et un nombre de centenaires si élevé. Après renseignement auprès du service des pompes funèbres, cette anomalie a été expliquée : quand le service ne connaît pas la date de naissance du défunt, il rentre soit 1900, soit la date du jour du décès. Le service des pompes funèbres était donc parfaitement au courant de cette anomalie et traitait donc ces données en ayant conscience de cela. A présent, le service a changé son mode de gestion pour éviter de nouvelles erreurs. Aujourd'hui, avec l'ouverture de ces données, ce genre de chose n'est plus possible, car il faut que la donnée publiée soit vérifiée et réutilisable par tous. Les services sont donc obligés de faire attention à la justesse de ce qu'ils produisent. Cela est un avantage lors de la mise en place d'outils décisionnels, mais aussi pour le travail des services au quotidien, puisque les réutilisateurs ont la garantie que la donnée publiée est juste et qu'ils ne perdent donc pas de temps à la vérifier.

De surcroît, de plus en plus d'agents de Toulouse Métropole partent à la retraite ces dernières années. Ces personnes avaient créé des réseaux pour récupérer des informations, ils savaient où aller les chercher. Comment procéder après leur départ en retraite ? L'open data permet en partie de pallier cette disparition de réseaux personnels grâce à la centralisation des données.

L'open data permet également de valoriser des ressources qui ne l'étaient pas avant. A titre d'exemple citons à Toulouse la bibliothèque municipale qui diffuse sur Flickr (site de partage de photographies et de vidéos en ligne) les clichés du photographe Eugène Trutat, qui datent du 19^e siècle.

Enfin, le développement d'applications comme « Qui Dit Miam ! » représente à coup sûr des économies pour la ville, puisqu'elle n'a pas eu à la développer. La start up, qui s'est servie des données publiées par la ville sur les cantines scolaires, emploie aujourd'hui 4 à 5 salariés

et l'estimation du développement de l'application s'élève aux alentours de 80 000 €. Finalement, c'est un service rendu gratuitement à la ville et aux ménages, grâce à l'ouverture de ces données.

L'étape suivante sera de croiser toutes ces données, afin de prendre du recul sur le travail effectué par la collectivité et d'observer si certains phénomènes, que personne n'avait observé avant ces croisements, prennent place dans la ville. Si tel est le cas, cela pourrait permettre à la métropole d'intervenir pour aider des dynamiques qui pourraient être positives pour la ville. Par exemple, s'il est observé un intérêt particulier pour le street art, à travers l'analyse de données liées aux tweets contenant des hashtags en lien avec le street art et Toulouse, ou une recrudescence d'évènements culturels d'associations en lien avec cet art, la ville pourrait proposer à son tour un évènement en lien avec cela : un festival, une exposition temporaire à l'échelle de la ville qui associe des artistes célèbres... autant de réponses possibles à un phénomène identifié par la donnée et qui représente un intérêt pour la population. Une plateforme big data, d'analyse des données est donc en train d'être élaborée. Ce nouvel outil sera interne mais pourra associer des partenaires de la collectivité. Il permettra de regrouper un volume important de données et de les croiser. Les données qu'il traitera ne seront pas toutes ouvertes, mais les données produites suite à des analyses de la plateforme pourront être libérées par la suite. Cet outil sera opérationnel d'ici un an.

La donnée ouverte favorise le développement économique et l'innovation :

Le développement économique et l'innovation sont des arguments qui ont été pointés lorsque la décision de se lancer dans l'open data a été prise, d'autant plus qu'une métropole innovante améliore son image et est donc plus attractive. L'ouverture des données a eu un impact positif au niveau économique pour la métropole, même si cela n'a pas été chiffré. Plusieurs entreprises se servent des données libérées sur la définition géographique des secteurs de Toulouse, pour calquer leurs activités sur ce découpage. L'avantage de ces découpages est qu'il fait gagner du temps aux entreprises dans la définition des secteurs où elles interviennent, d'autant plus que ces découpages sont officiels et utilisés par la ville, ce qui fait qu'ils ne sont pas contestables. Des entreprises informatiques récupèrent aussi des

jeux de données pour tester les outils qu'elles développent, ce qui leur permet de réaliser des économies en évitant d'en réaliser elles-mêmes.

Chiffrer les retombées économiques liées à l'ouverture de la donnée est toutefois complexe, d'autant plus qu'aucun état des lieux n'a été dressé au préalable. La difficulté à évaluer cet impact est due au fait que la collectivité n'a pas forcément de retour sur ce qui a été réutilisé et à quelles fins. Les entreprises qui réutilisent les données sont gagnantes mais elles ne chiffrent pas forcément l'économie (de temps et financière) qu'elles réalisent grâce à ces réutilisations.

Quelques applications concrètes de la donnée ouverte à Toulouse :

- ***Un concours pour animer la réutilisation de la donnée ouverte :***

L'innovation peut être favorisée par des concours en lien avec la réutilisation de données ouvertes. A Toulouse par exemple, le concours « Défi numérique » a permis d'inciter au développement d'applications entre 2012 et 2014³⁴. Certaines applications favorisent le tourisme, comme « DOT » qui propose des itinéraires adaptés à un point de départ et au temps disponible pour visiter, ou encore Flora'um, application participative qui permet de découvrir la flore de Toulouse ou d'autres villes.

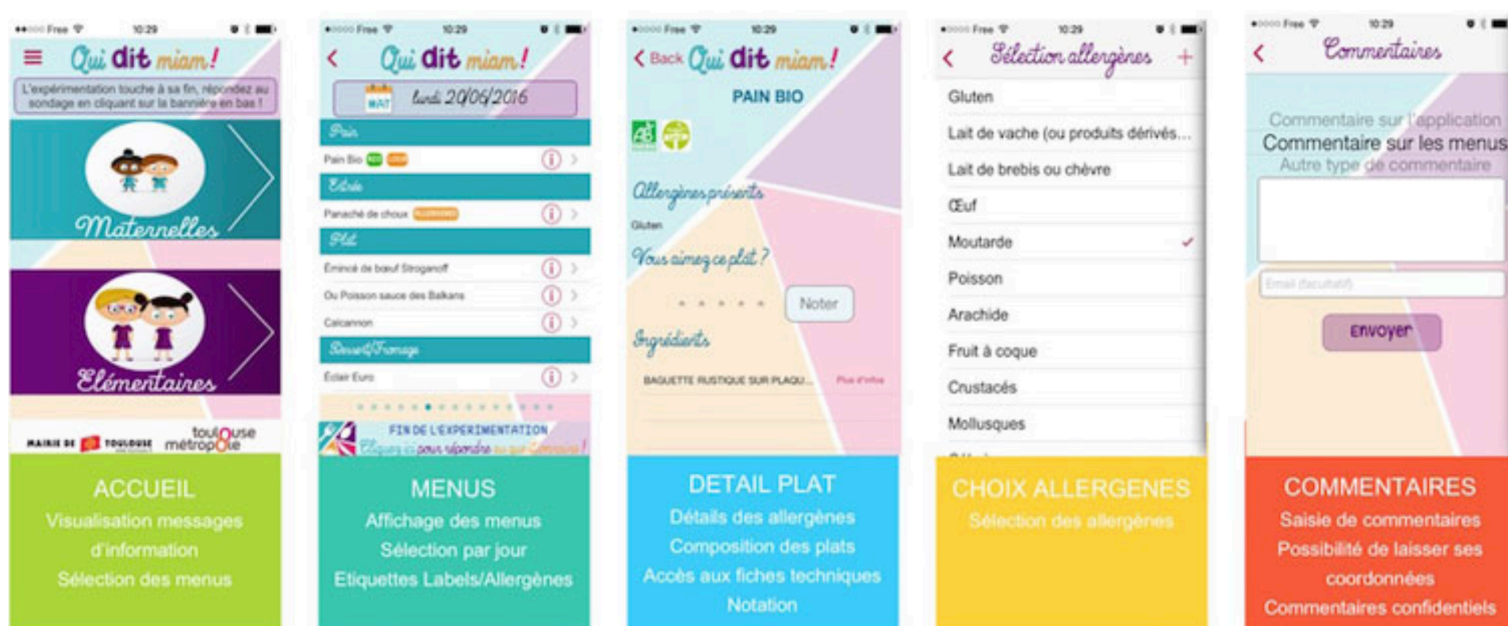
- ***Une multitude d'applications qui rendent service aux citoyens dans leur vie quotidienne :***

Avec la participation du réseau de transports toulousain Tisséo, plusieurs applications ont pu être créées en lien avec la mobilité, comme « Bus 31 ». Développée en 2013 par un particulier, elle permettait de connaître les horaires de bus en temps réel, comme celle développée par Tisséo aujourd'hui. L'application « Allô Toulouse » utilise les données

³⁴ Toulouse Métropole. *Concours défi numérique : défiez les données ouvertes par Toulouse Métropole* [en ligne]. (2014) Disponible sur : < https://www.toulouse-metropole.fr/projets/les-projets/-/asset_publisher/VCxskVEr4zgQ/content/concours-defi-numerique-defiez-les-donnees-ouvertes-par-toulouse-metropole;jsessionid=E6B4D84B77E04F58E202FAA9CD4777BC > (consulté le 01/06/18)

ouvertes, elle permet aux habitants de signaler des nuisances liées au bruit, aux encombrants, aux conflits de voisinage etc. Pour la ville, c'est un moyen de quantifier ces nuisances et de les centraliser afin de les traiter plus rapidement, sans encombrer les services de police. Pour connaître l'agenda des manifestations sur Toulouse, l'application Urban Pulse permet de géolocaliser les événements et de préciser leur contenu. Avec l'application « Qui Dit Miam ! », développée par la startup GOA SOFT, les parents des enfants scolarisés à Toulouse peuvent connaître le menu de la cantine pour l'ensemble de la semaine, mais aussi la liste des allergènes par repas. Les parents peuvent aussi noter les repas de leurs enfants et reçoivent des notifications pour signaler les mouvements de grèves des cantines. L'application totalise à ce jour plus de 5 000 téléchargements, et tout cela est possible grâce à l'ouverture des données sur les repas de la cuisine centrale de la ville.

Illustration 8 : l'application "Qui Dit Miam !" (source : <https://www.opendatasoft.fr/2016/10/17/y-a-d-frites-application-avec-de-lopen-data-dedans/>)



Depuis peu, la géolocalisation des espèces végétales en ville permet d'identifier les zones à risques pour l'émission de pollen, mais aussi de savoir quel type de pollen est susceptible d'être présent dans l'air et à quel moment de l'année. Cela peut faciliter la vie des personnes allergiques qui, grâce à une simple application smartphone (Météo Pollen), peuvent éviter les zones à risque. Enfin, une application sur la ville de Toulouse, développée par la ville, est en cours de réalisation avec une sortie prévue en mars 2019. Elle reprendra l'agenda de la ville (notamment activités sportives, culturelles...), intégrera une partie d'Allô Toulouse mais permettra également d'identifier toutes les applications existantes qui concernent Toulouse (développées par la collectivité ou le privé).

Illustration 9 : Météo Pollen, une application créée par une startup toulousaine (source : Météo Pollen, une nouvelle application lancée par une start-up toulousaine (source : <https://meteopollen.com>)

MÉTÉO POLLEN
☰

POLLEN(S)

Une manière simple et rapide d'être informé sur la présence de pollens, partout en France métropolitaine.

Source :  Atmosphere Monitoring Service
atmosphere.copernicus.eu

GÉOLOCALISATION

Des informations adaptées à votre position géographique.

MÉTÉO

Retrouvez en un coup d'oeil les informations météorologiques importantes.

Source :  Météo France

PRÉVISION À 2 JOURS

Météo Pollen vous informe sur les niveaux de pollution et les prévisions de la qualité de l'air dans votre environnement, pour les 2 prochains jours.

Source :  Atmosphere Monitoring Service
atmosphere.copernicus.eu

QUALITÉ DE L'AIR

Air sain ? Air pollué ? Retrouvez l'indice de qualité de l'air, là où vous vous trouvez.

CONSEILS

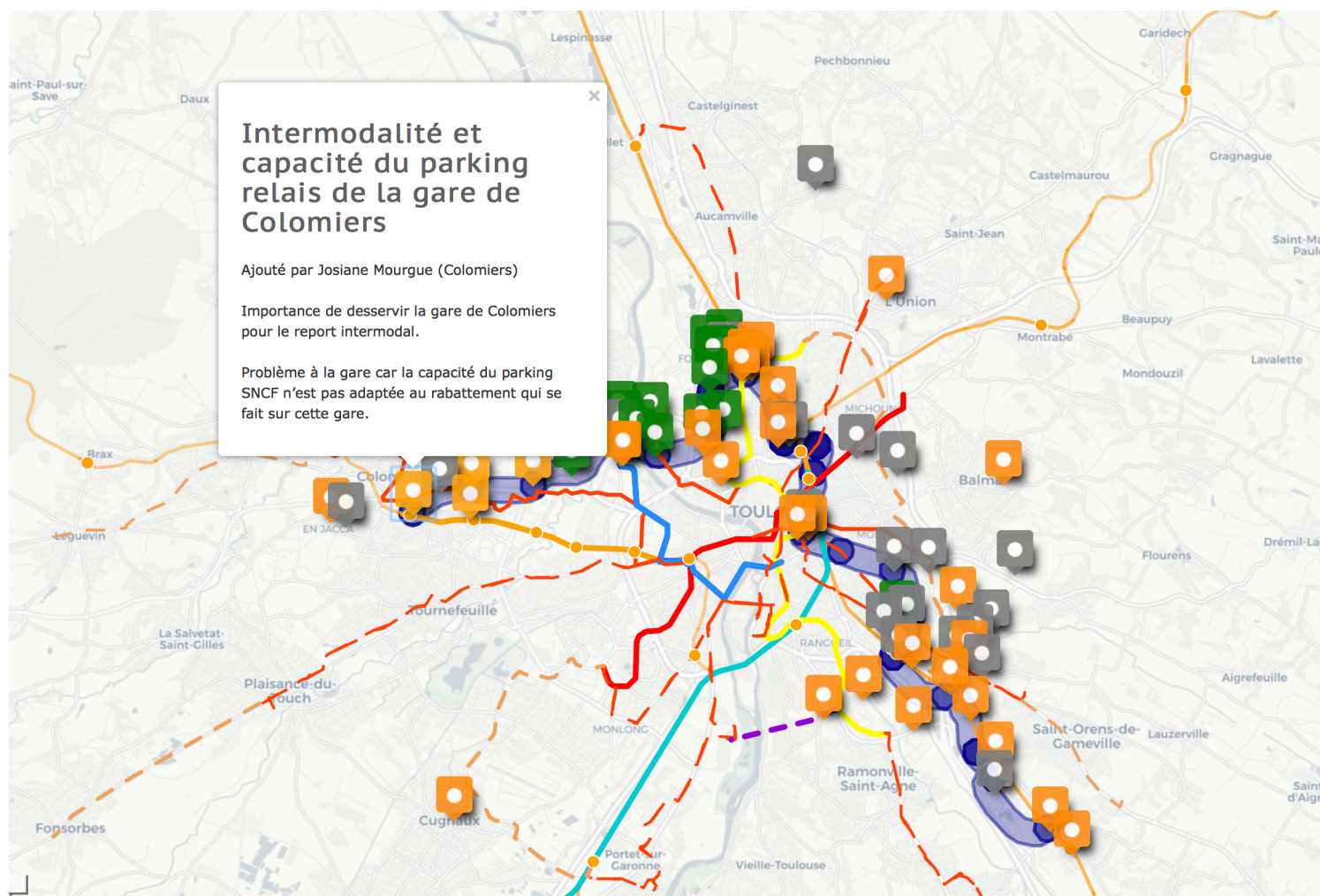
Cette section vous permettra d'en savoir plus sur les pollens et d'adopter les gestes pour vous prémunir des allergies.



- **Les données peuvent aussi servir au débat public :**

Pour la phase de débat public qui concerne la 3^{ème} ligne de métro à Toulouse, des données ouvertes ont permis aux participants de réaliser des cartes pour s'exprimer autour du projet. Celles-ci ont été regroupées dans une cartothèque sur le site du débat (metroligne3toulouse.debatpublic.fr) et ont servi à la discussion. Une carte interactive des enjeux locaux a également été produite : chacun pouvait y inscrire des remarques sur le projet. De même, dans le cadre du projet de renouvellement urbain du quartier Trois Cocus – La Vache, la population est invitée à s'exprimer sur un registre en ligne, accessible à tous depuis le site de Toulouse Métropole (toulouse-metropole.fr/projets).

Illustration 10 : carte interactive des enjeux locaux, produite dans le cadre du débat de la 3^{ème} ligne de métro (source : https://framacarte.org/fr/map/carte-des-enjeux-locaux_5330#12/43.6200/1.4396)



3. ZOOM SUR LE PHENOMENE OPEN DATA : UN OUTIL STRATEGIQUE POUR LES COLLECTIVITES

Illustration 11 : nuage de mots (réalisé avec : nuagesdemots.fr)



Le mouvement open data s'inscrit donc, comme cela a été exposé précédemment, comme une tendance, accompagnée d'une volonté politique bien présente, qui accompagne le phénomène à travers des textes de lois qui se multiplient au fil des ans. Derrière l'engagement des collectivités vers l'open data, il y a donc une obligation légale, mais ce n'est pas tout. Dans cette dernière partie nous nous intéresserons aux nombreux avantages liés à l'ouverture de la donnée. L'avantage économique est particulièrement important, d'autant plus que le potentiel des économies à réaliser grâce à l'open data est non négligeable, compte tenu de la conjoncture économique actuelle et des restrictions budgétaires de l'Etat.

3.1. La donnée ouverte comme bien commun, favorisant l'innovation économique et sociale

Pour les acteurs publics : plus d'efficacité, de transversalité et d'économies à la clé

L'ouverture des données peut permettre aux collectivités de se moderniser et de gagner en efficacité : centralisation des données, pas de perte de temps à vérifier la véracité des jeux diffusés... L'open data apporte également plus de transversalité entre les services, ce qui à terme permet souvent la production de nouvelles connaissances, tout cela grâce à la transformation de la donnée en un bien commun. D'autre part, la qualité de la donnée s'améliore, puisque les services font plus attention à ce qu'ils publient : chaque jeu doit être sûr et réutilisable par tous avant d'être ouvert. Autre bénéfice inattendu de la donnée : le remplacement des réseaux d'employés qui s'étaient créés au fil du temps pour récolter de la donnée, mais qui sont en péril lors des vagues de départs à la retraite dans les administrations. L'échange de données entre les services permet aussi d'éviter les redondances, en recréant des bases de données qui existent déjà dans d'autres services.

Aujourd'hui toute la donnée est centralisée sur une plateforme, ce qui fait en plus gagner du temps car la donnée est déjà vérifiée. D'ailleurs, les économies réalisées par les collectivités ne sont pas négligeables : économies sur les achats de données puisqu'elles sont aujourd'hui gratuites, économies de main d'œuvre (gain de temps), économies liées au développement d'applications (créées par le privé grâce à la réutilisation de données ouvertes). A Washington, un concours lancé par la ville en 2010 pour créer des applications à partir de données publiques ouvertes a permis la création de 47 applications en 30 jours³⁵. La ville de Washington a investi 50 000 \$ dans ce concours, mais le développement de toutes ces applications lui aurait coûté 2 millions de \$, et certainement beaucoup plus de temps. C'est donc un gain important pour la ville, sans compter la création d'emplois qui est générée par cette activité.

³⁵ Better Government Competition. *Apps for Democracy* [en ligne]. (19/07/10) Disponible sur : < <http://bgc.pioneerinstitute.org/apps-for-democracy/> > (consulté le 05/06/18)

Enfin, l'open data permet de mieux évaluer l'action publique. Au lieu d'évaluer les effets d'une politique par un seul service, le partage des données permet à d'autres services de contrôler l'impact que cela peut avoir sur leur domaine. Par exemple, les politiques qui concernent le logement intègrent bien souvent aujourd'hui l'enjeu environnemental : grâce au partage des données, le service environnement peut évaluer les effets des actions menées par le service habitat. De même, les dépenses publiques peuvent être suivies. Au Nigéria, le site « Follow the Money » permet aux citoyens de vérifier que les fonds publics sont bien investis dans les politiques menées par l'Etat au service de la population³⁶, réduisant ainsi les risques de corruption, mais aussi de « mauvais investissements » de l'argent public.

Des enjeux économiques non négligeables, mais aussi environnementaux, doublés par l'innovation :

L'ouverture de la donnée offre également une opportunité de création de nouvelles entreprises. Avec le développement d'applications, plusieurs startups voient le jour : Météo Pollen ou encore Global Biker (qui a créé la première communauté de voyageurs à vélo) en sont des exemples pour la ville de Toulouse, mais cela se multiplie partout à travers le monde et peut parfois donner naissance à des entreprises plus importantes (comme le GPS communautaire Waze). Au Danemark, l'ouverture des données postales, auparavant payantes, a permis la création de 48 entreprises et de 90 emplois.

A retenir aussi, l'effet de marketing territorial que l'open data peut avoir, favorisant à terme l'installation d'entreprises. En effet, l'ouverture des données permet à une entreprise d'observer le tissu économique local et de réfléchir en amont à ses possibilités d'installation. Cela est d'autant plus attractif si de nombreuses entreprises déjà installées réutilisent des données publiques ouvertes et repartagent leurs utilisations, ce qui démontre le dynamisme économique d'un territoire.

³⁶ European Commission. *L'open data, c'est quoi ?* [en ligne]. (2018) Disponible sur : < <https://www.europeandataportal.eu/elearning/fr/module1/#/id/co-01> > (consulté le 10/05/18)

Grâce aux jeux de données disponibles les entreprises peuvent tester leurs nouveaux outils à moindre coût, puisqu'elles n'ont plus besoin d'en créer. Elles économisent du temps et donc de l'argent sur la masse salariale, qui peut être réorientée vers d'autres ambitions. Avec l'open data, tout comme dans l'administration publique, une meilleure transversalité s'inscrit dans le tissu économique des territoires. Les acteurs privés peuvent superposer et combiner ces nouvelles données, se les approprier pour créer de nouvelles informations et donc de nouvelles connaissances qui donneront lieu à de nouveaux usages et donc à de l'innovation.

Finalement, pour le territoire, l'open data peut permettre de dynamiser le tissu économique existant, de l'enrichir par l'accueil de nouvelles entreprises et l'innovation qui peut en résulter voire participer à la réduction du taux de chômage sur un territoire. De même, des ménages, et donc des actifs, qui souhaitent s'installer quelque part peuvent se renseigner facilement sur leur futur lieu de vie en termes de service, d'aménités, de culture etc. La ville est d'autant plus accueillante qu'une forme de modernité est apportée par la mise à en ligne de ces données, mais aussi une forme de communauté impliquée dans sa ville et qui a envie de partager et donc possiblement d'accueillir.

L'open data peut être également bénéfique à l'environnement : en libérant de la donnée sur l'eau, le climat, la biodiversité... Les jeux de données peuvent être récupérés et réutilisés afin de produire de nouvelles connaissances et de nouveaux outils, au service de la protection de l'environnement et de la gestion des risques : prévention des risques d'inondation, protection de la biodiversité, applications pour une meilleure gestion de l'eau... La simple libération des jeux de données liés aux transports en commun permet une meilleure visibilité du service pour les usagers et peut susciter l'augmentation de leur fréquentation, ce qui peut influencer alors sur la réduction des émissions de gaz à effet de serre dans les villes.

Enfin, la recherche en sciences sociales peut elle aussi être aidée par l'open data, qui va faciliter l'accès à la donnée pour les chercheurs : « la masse de donnée doit permettre de faire émerger des régularités, mais aussi de trouver des lois du fonctionnement du monde

social »³⁷. Mais certains scientifiques, comme Etienne Ollion, pointent aussi les précautions à prendre avant l'utilisation de cette data, dont notamment la manipulation de masses de données, qui peuvent rassembler l'ensemble d'une population, mais pour autant ne constituent pas forcément un échantillon aussi qualitatif que ceux réalisés auparavant par les chercheurs (sans oublier les marges d'erreurs possibles dans ces fichiers). De plus, il faut également que les chercheurs aient les compétences techniques, en matière d'informatique, pour récupérer et traiter cette donnée. Pour cela, les chercheurs en sciences sociales font appel à leurs homologues en sciences de l'informatique, ce qui conduit finalement à travailler de manière interdisciplinaire par le biais de l'utilisation de la data et favorise donc les échanges entre différents domaines scientifiques.

L'open data comme outil de la participation citoyenne :

Les contraintes financières des collectivités, couplées de l'émergence de nouveaux défis tels que le vieillissement de la population, la prise en compte de l'écologie, la mixité urbaine, la démocratie participative... font que l'action publique ne peut plus être la seule à répondre à ces nouveaux enjeux. Grâce à l'ouverture des données, elle peut déléguer des tâches : de la réflexion au développement de projets qui répondent aux enjeux de demain, tout en favorisant des solutions innovantes. De fait, il n'y a plus seulement des acteurs de la sphère privée qui peuvent intervenir, mais aussi des citoyens, des associations... L'ensemble des acteurs peut prendre part au débat public.

Aujourd'hui, l'idée est de faire en sorte qu'à travers le nouveau bien commun qu'est la donnée, l'innovation publique, sociale et économique se rejoignent afin d'améliorer le fonctionnement des territoires.

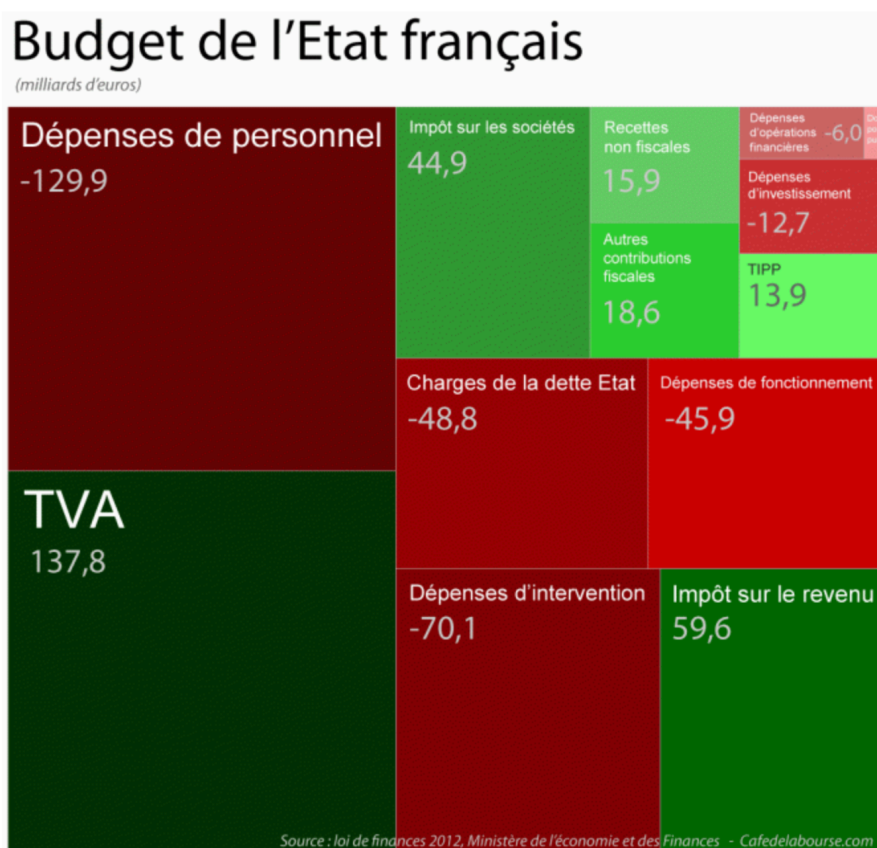
Libérer de la donnée permet de toucher plus de citoyens. Ils disposent des mêmes outils de compréhension d'un projet et peuvent donc s'exprimer de manière plus égalitaire, grâce à une information partagée et accessible par tous. L'intervention de citoyens en ligne peut

³⁷ OLLION E. L'abondance et ses revers. Big Data, open data et recherches sur les questions sociales [en ligne]. *Informations sociales*, 2015, n°191, pages 70 à 79. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-70-.htm> > (consulté le 19/06/18)

favoriser la prise de parole d'individus aux profils plus variés que ceux que l'on retrouve dans les débats publics, qui se ressemblent et représentent un public plus averti qui parfois monopolise la parole, limitant l'intervention de publics moins aguerris.

L'open data permet également une meilleure lisibilité de l'action publique. Pour faciliter l'accès aux données au plus grand nombre, il faut privilégier les outils de visualisation de la donnée en ligne avec par exemple des cartes interactives. La data visualisation (la représentation de données de manière visuelle) est un bon outil pour vulgariser des données,

Illustration 12 : exemple de treemap - Le budget de l'Etat français en 2012
(source : www.cafedelabourse.com)



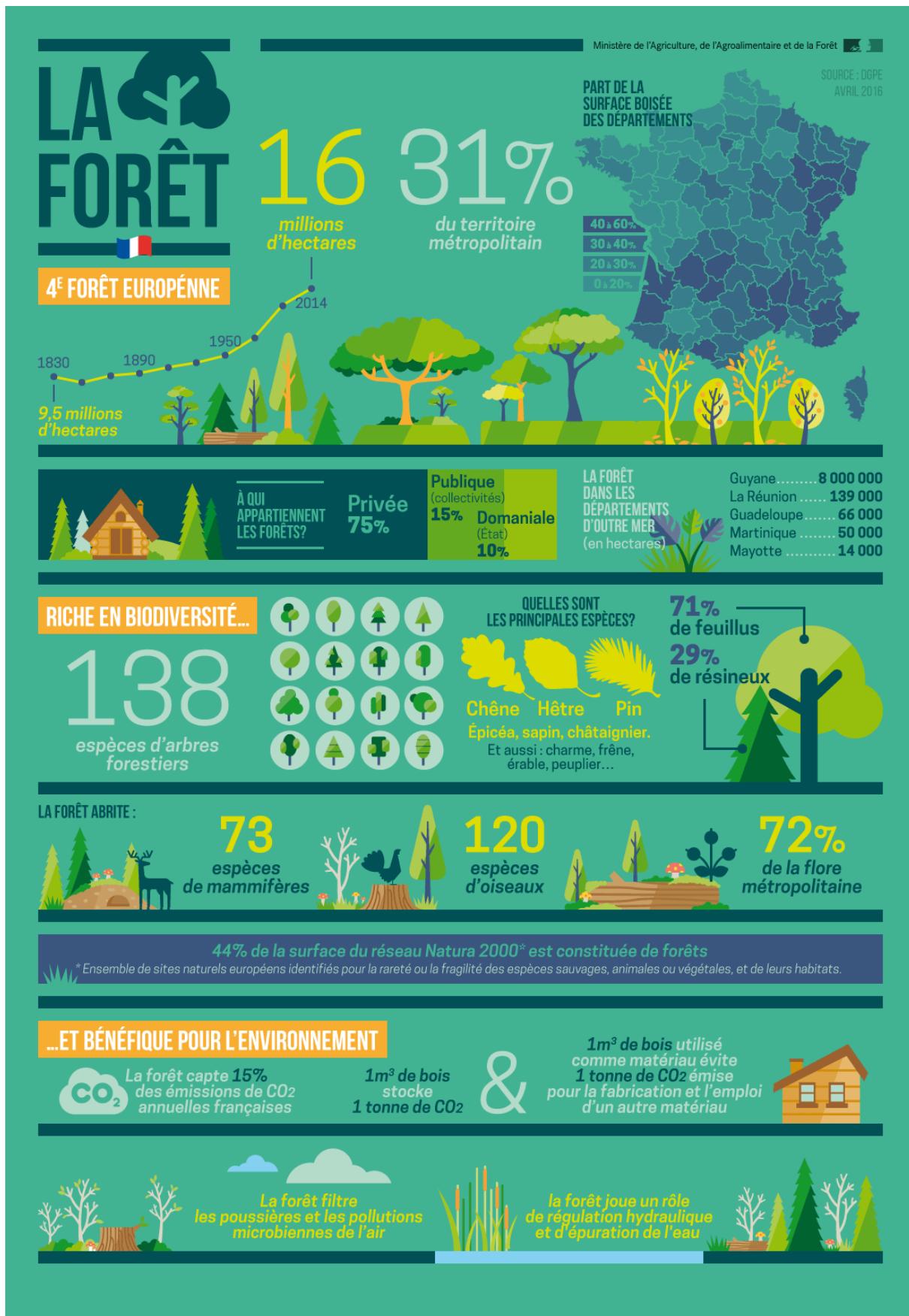
tout comme le treemapping (qui représente les données de manière hiérarchique).

En ouvrant la donnée, les collectivités prônent la transparence.

L'échange d'informations renforce leur légitimité en rendant plus compréhensible l'action publique, ce qui participe à sa valorisation. Ainsi, un

meilleur dialogue s'installe entre administrateur et administrés, ce qui favorise l'implication des citoyens dans la vie de la commune et peut aboutir à terme à de la co-production sur des projets menés par la collectivité. L'open data permet donc d'aboutir à une forme d'*empowerment* des citoyens : ils peuvent plus facilement agir et s'organiser.

Illustration 13 : infographie - la forêt française (source : agriculture.gouv.fr)



En Haïti, après le tremblement de terre de 2010, le cartographie du pays a été libérée par l'Etat sur OpenStreetMap, pour que la communauté enrichisse et corrige les cartes. Cette action a permis de récolter gratuitement de nouvelles informations après la catastrophe naturelle : accessibilité aux personnes handicapées, points d'intérêt, nouveaux tracés des rues...

En France, en 2009, la commune de Plouarzel (Finistère) a fait participer la population locale à la mise à jour du plan communal. Face au manque de moyens, la commune a été inventive. En partenariat avec le centre multimédia de Plouarzel, les habitants ont été équipés de GPS et ont cartographié les sites remarquables de leur commune, à l'occasion d'une balade à vélo, en voiture, à pied ou encore en kayak. Les données récoltées sont aujourd'hui utilisées par l'Office du tourisme. En plus de réaliser des économies pour la commune, cette opération a eu un réel avantage social : avec cette méthode innovante, chaque citoyen a pu rendre directement service à la communauté. Cette opération a permis de favoriser les échanges au sein de la population, de leurs centres d'intérêt mais aussi de leurs connaissances.

Lorsque la donnée n'est pas partagée d'abord par les acteurs publics, ils prennent le risque de faire apparaître des soupçons de dissimulation : en libérant un jeu de données sur l'attribution de subventions par exemple, une collectivité peut lever le doute sur des soupçons de corruption ou de mauvaise utilisation de l'argent public. Les informations non délivrées risquent d'être quand même divulguées par des groupes qui ont un intérêt contraire aux collectivités et qui peuvent chercher à leur nuire, d'où l'intérêt de prendre les devants³⁸. L'open data est aussi une manière d'évaluer autrement l'efficacité de l'action publique et d'identifier les possibles dysfonctionnements. L'association Regards Citoyens a créé la plateforme www.nosdeputes.fr : elle permet de partager les activités des députés (absences, travaux en cours etc.)³⁹. Cette plateforme peut pointer les erreurs commises par les députés, mais elle a aussi un rôle de mise en valeur du travail politique car elle permet à chaque internaute de s'informer, mais aussi de s'exprimer et pourquoi pas de participer aux réflexions à travers les commentaires qu'ils peuvent laisser.

³⁸ FING. *Guide pratique de l'ouverture des données publiques territoriales* [en ligne]. (janvier 2011) Disponible sur : < <http://doc.openfing.org/RDPU/GuidePratiqueDonneesPubliquesv1beta.pdf> > (consulté le 04/05/18)

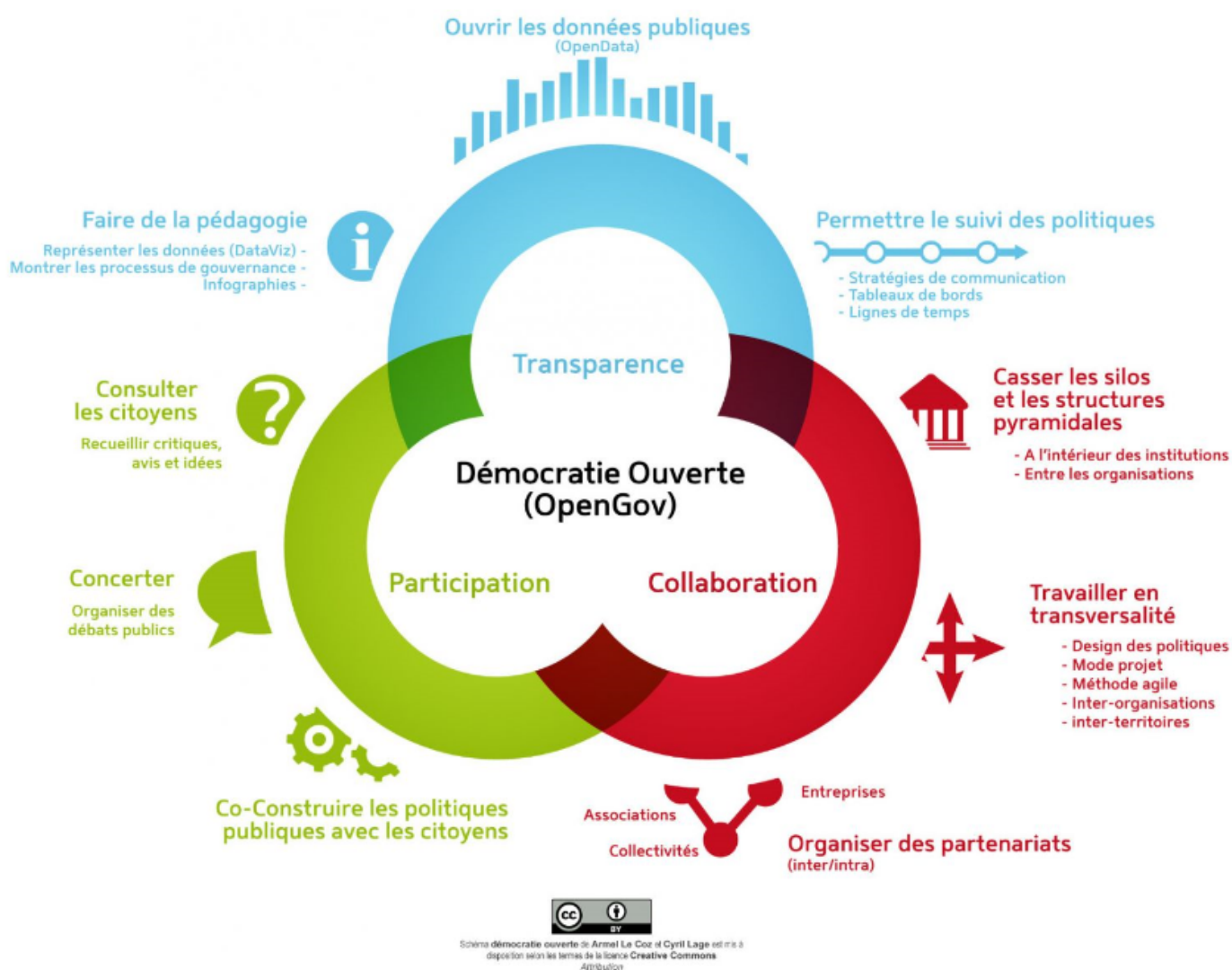
³⁹ CHIGNARD S. *Comprendre l'ouverture des données publiques*. Quercy : éditions FYP, 2012, 191 p.

Illustration 14 : le site de l'association Regards Citoyens (source : nosdeputes.fr)



En résumé, l'open data offre de nouvelles opportunités d'accès à des informations et de dialogue entre sphères d'acteurs, ce qui peut séduire les collectivités, mais également intéresser la sphère privée et la société civile. Ce phénomène concerne une multitude d'acteurs et les incite à prendre part et à s'informer des décisions politiques. L'ouverture de la donnée offre la possibilité à chacun de devenir acteur de son territoire et d'améliorer la qualité, le cadre de vie des habitants de son quartier, de sa ville, voire de territoires plus lointains.

Illustration 15 : Logiques et valeurs applicables à l'open data (extrait du site : opendatalab.fr)



3.2. L'open data en France et à l'étranger

En France et en Europe, illustration de l'utilisation de l'open data pour dynamiser l'économie, maîtriser les flux et protéger l'environnement :

En France, depuis 2014, la ville de Paris a intégré une « clause open data » dans ses marchés publics, afin d'obliger les prestataires à ouvrir leurs données pour qu'elles puissent être réutilisées. La ville mise sur le développement de start-ups et l'innovation numérique. Depuis le lancement de sa politique open data, 100 000 m2 ont été dédiés à l'aménagement

de lieux pour accueillir des start-ups, dont 3 000 m² sous la Halle Freyssinet, qui accueille aujourd'hui une multitude d'entrepreneurs.

A Nice, le boulevard Victor Hugo est devenu un boulevard connecté : il est équipé de 200 capteurs dans des lampadaires, la chaussée et des bennes à ordures. Les informations récoltées permettent aux automobilistes de trouver plus facilement une place de parking, de faire fonctionner les lampadaires seulement lorsqu'il y a présence de piétons ou encore de mesurer la pollution atmosphérique⁴⁰.

La métropole lyonnaise a développé « Optimod' »⁴¹, avec l'aide de plusieurs partenaires : cette application rassemble un ensemble de données sur le trafic et les transports, en temps réel, afin de proposer les meilleurs itinéraires aux utilisateurs, en combinant tous les modes de transport disponibles dans la ville. L'application informe aussi des retards et autres problèmes sur les lignes et propose des solutions alternatives en cas de panne.

Plusieurs villes européennes se sont aussi lancées dans l'ouverture de leurs données. Aux Pays-Bas, la ville d'Amsterdam est connue pour ses politiques en faveur de la protection de l'environnement. La ville anticipe que sa population va augmenter dans les prochaines années, et ses dépenses énergétiques avec. Cependant elle souhaite maîtriser ces dépenses et réduire de 40% les émissions de CO₂ (par rapport au niveau de 1990) d'ici à 2025⁴². La ville a donc choisi d'ouvrir les données qui concernent la production et la consommation d'énergie.

Associée à des bailleurs et des entreprises privées, la municipalité a créé un atlas de l'énergie qui rassemble les données relatives à la consommation énergétique à l'échelle du quartier. Grâce à cet outil, il est aujourd'hui possible d'identifier les zones entre lesquelles des

⁴⁰ NODE-LANGLOIS F., EGLOFF E. Nice inaugure on « boulevard connecté » [en ligne]. LeFigaro.fr (18/06/13) Disponible sur : <<http://www.lefigaro.fr/conjoncture/2013/06/17/20002-20130617ARTFIG00709-nice-inaugure-son-boulevard-connecte.php>> (consulté le 15/06/18)

⁴¹ Application Optimod' Lyon. Disponible sur le Google Play Store : <https://play.google.com/store/apps/details?id=fr.cityway.android_v3.optymod&hl=fr>

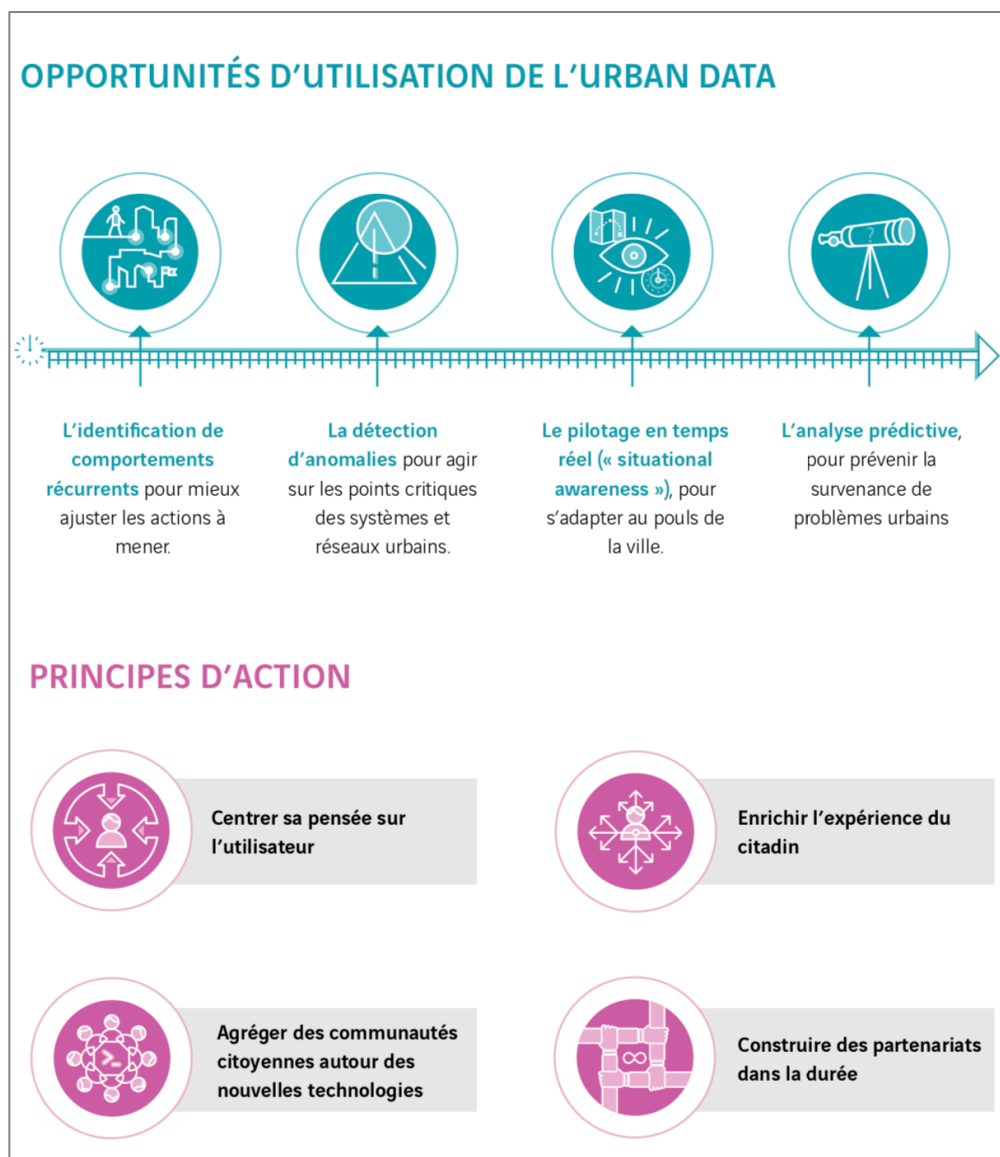
⁴² Thibaud (OpenDataSoft). *L'Energy Atlas d'Amsterdam : vers la ville durable grâce à l'urban data* [en ligne]. (22/02/17) Disponible sur : <<https://www.opendatasoft.fr/2017/02/22/lenergy-atlas-damsterdam-vers-la-ville-durable-grace-a-lurban-data/>> (consulté le 06/06/18)

échanges d'énergie peuvent avoir lieu. Cet outil permet de sensibiliser les usagers à leurs dépenses énergétiques, mais aussi de découvrir les potentiels liés aux énergies vertes pour la ville. Plusieurs cartes peuvent être consultées : la déperdition de chaleur par les toits, la carte des potentiels solaires pour identifier les endroits où implanter des panneaux solaires, mais aussi la carte des déchets pour repérer les zones qui en produisent le plus.

En Allemagne les villes de Hambourg, Berlin ou encore Francfort, disposent d'une application pour géolocaliser des terrains et des bâtiments abandonnés. L'application est enrichie par les utilisateurs. C'est un outil gratuit à disposition de la collectivité, qui peut

Illustration 16 : les utilisations possibles de l'open data (source : lafabriquedelacite.com)

identifier des lieux à développer. Mais c'est également une incitation à occuper temporairement ces lieux abandonnés, par des initiatives citoyennes, avec la création de lieux culturels ou de jardins partagés.



Crowdsourcing⁴³ et participation citoyenne à Boston :

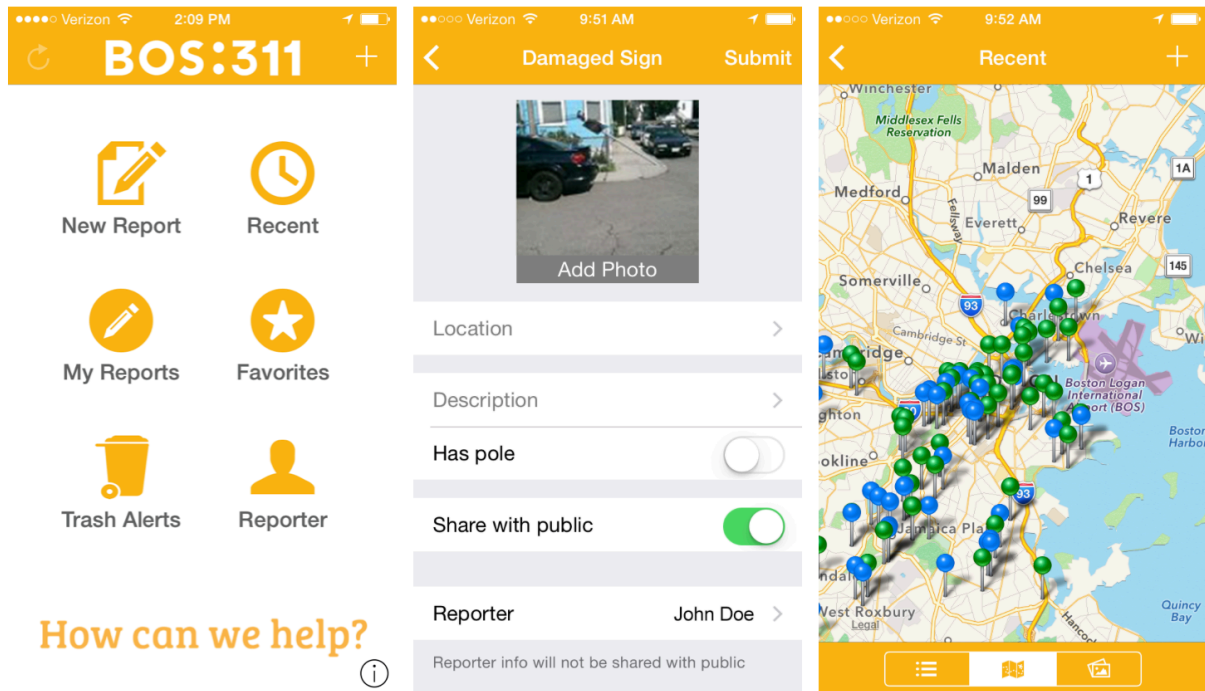
La ville de Boston a centré ses efforts en matière d'open data sur des applications qui améliorent le quotidien des citoyens. Par exemple, pour signaler des problèmes dans la ville, la municipalité propose d'appeler le numéro 311, mais pour améliorer ce service, l'application Citizen's Connect est créée par la ville en 2009. Grâce à son smartphone, tout citoyen peut signaler un problème, mais aussi le prendre en photo et le géolocaliser : tout cela est envoyé instantanément aux services urbains, ce qui facilite la résolution des problèmes signalés⁴⁴. Mais l'application va plus loin, car elle permet à l'utilisateur de voir si le problème a été résolu par la ville, mais aussi de connaître l'identité des agents qui s'en sont occupés et de les féliciter, ou de les dénoncer... Les utilisateurs peuvent aussi obtenir le classement des « meilleurs citoyens », ce qui les incite à participer au projet et à s'impliquer dans la vie de leur ville. La ville a toutefois observé que le nombre de signalements avait quasiment doublé, mais ce n'est pas un souci, car la grande majorité sont traités par la ville. Cette application s'est montrée d'une grande aide après le passage de l'ouragan Sandy sur la ville en 2012 : combinée à une application utilisée par les agents municipaux pour traiter et planifier leurs actions, les citoyens ont pu relever les incidents observés en temps réel et les envoyer directement aux services concernés. Pour impliquer encore plus les citoyens dans la ville, un portail nommé « Boston About Results » a été créé : il permet d'observer les résultats de politiques publiques menées par la ville en lien avec par exemple le taux de chômage, l'ouverture d'entreprises etc. Plusieurs autres applications permettent de mieux vivre en ville au quotidien : « Where's My School Bus » est utilisé par les écoliers et leurs parents pour connaître en temps réel le déplacement des bus scolaires (très utile en hiver pour anticiper les problèmes liés à la neige

⁴³ **Crowdsourcing** : « *recueil ou création de données par un groupe de volontaires, qui aboutit à la constitution d'une base de données participative. Le crowdsourcing peut s'appliquer à de multiples secteurs d'activité (urbanisme, création artistique, design...).* » (source : cf. référence 44)

⁴⁴ La Fabrique de la Cité. *Comment optimiser les villes avec les data ?* [en ligne]. (23/03/2015) Disponible sur : < [https://www.lafabriquedelacite.com/fabrique-de-la-cite/data.nsf/951F68AEB0B75B1CC1257E0F003225D2/\\$file/etude_def_urbandata2.pdf](https://www.lafabriquedelacite.com/fabrique-de-la-cite/data.nsf/951F68AEB0B75B1CC1257E0F003225D2/$file/etude_def_urbandata2.pdf) > (consulté le 02/06/18)

et aux retards de bus) ; « Street Bump » est une application collaborative qui permet de signaler les nids de poule sur les routes grâce aux capteurs de mouvements des smartphones des utilisateurs (accéléromètre, inclinaison, géolocalisation...).

Illustration 17 : l'application citizens connect à Boston (source : boston.gov)



Pour tenter de comprendre des phénomènes urbains, les données peuvent être croisées pour trouver des réponses. A Boston un lien a été établi entre le nombre de vitres cassées et les difficultés sociales de certaines zones, comme l'échec scolaire, le trafic de drogue, la violence... Il est alors possible de se questionner sur le traitement des problèmes identifiés et les risques de stigmatisation entraînés par ce genre de croisement.

Une ville plus vivable et une police prédictive à Los Angeles :

Afin de réduire les embouteillages en ville, causés notamment par les automobilistes qui cherchent des places de parking, la municipalité de Los Angeles s'est associée à l'entreprise « XerosxLabs ». Des capteurs ont été placés sur les routes dans le quartier d'Hollywood qui était particulièrement touché par les embouteillages. Grâce aux données récoltées, les prix des places de parking sont ajustés selon la demande et la disponibilité des places de stationnement. Quant aux automobilistes qui ne trouvent pas de place de parking, ils peuvent utiliser l'application « Parker ! » qui localise les places de stationnement libres. Ces deux outils auraient permis de réduire le trafic automobile de presque 10% dans la ville, incitant les habitants à prendre les transports en commun et réduisant les émissions de gaz à effet de serre.

Encore plus innovant : depuis 2013 la police de Los Angeles utilise l'algorithme de la société Predpol, qui permet de prédire où et quand peut avoir lieu un délit dans la ville, en combinant des données spatiales et temporelles croisées avec la nature et la récurrence des délits commis sur une zone donnée. Cet algorithme permet aux policiers de réaliser des patrouilles pour dissuader les auteurs de possibles délits en ville, en ciblant une zone à 50 mètres près. Dans un quartier de Los Angeles qui a expérimenté cette méthode, le nombre de délits commis a baissé de 20% en un an. Aujourd'hui la police prédictive s'est développée dans d'autres villes américaines, mais aussi au Royaume-Uni.

D'autres initiatives ont été lancées, au niveau de la santé publique par exemple : sur l'application « Yelp » (qui permet de géolocaliser des restaurants et autres services notés par les utilisateurs), un indicateur sur l'équilibre alimentaire des repas proposés a été rajouté grâce aux données récoltées par le service d'hygiène de la ville. Le « City's Performance Dashboard » permet aux habitants d'observer l'efficacité des politiques publiques de la ville en leur fournissant des chiffres sur l'évolution du taux de chômage, la proportion d'électricité verte produite... associés aux actions menées par la municipalité pour tenter d'agir sur ces thèmes. Pour prolonger la volonté de transparence et d'accessibilité la ville diffuse également son budget sur un portail.

A New York, des plateformes open data fournies et efficaces :

De nombreuses applications ont été développées grâce au concours New York City Big Apps, organisé par la ville depuis 2009. « Heat Steek » est une application qui fonctionne avec des capteurs thermiques installés dans des appartements à la demande de l'utilisateur. Les habitants peuvent recevoir des conseils pour économiser de l'énergie et peuvent signaler les appartements qui ne respectent pas les normes d'isolation thermique.

Le portail Open Data de la ville de New York est la plus grande plateforme open data au monde : plus de 1 600 jeux de données sont ouverts⁴⁵ et ont été consultés plus de 2,8 millions de fois sur la première année de mise en ligne. La ville dispose aussi d'une plateforme open data similaire en interne, « NYC DataBridge ». Celle-ci regroupe également des données qui proviennent d'opérateurs publics (fournisseurs d'électricité par exemple) et d'universités. Les services urbains sont donc améliorés par la transversalité que permet ce partage de données, mais ont également plus de facilité à trouver les interlocuteurs privilégiés pour chacune de leur action. Cette plateforme a permis entre autre : d'identifier des pharmacies qui trafiquent des ordonnances, de détecter plus facilement la fraude hypothécaire... Mais aussi de lutter contre les incendies domestiques, car la ville manquant de logements, certains propriétaires divisent illégalement leur appartement ou utilisent des locaux commerciaux pour loger des ménages. Cela est possible grâce au croisement de données sur les signalements de modifications illégales d'appartements avec les plaintes déposées au 311 et d'autres informations (raccordements au réseau, non-paiement des loyers...), qui orientent au mieux les services d'inspection des risques incendies : cette réutilisation de données a permis de faire passer le taux de détection des infractions majeures de 8 à 70% des inspections.

La ville de New York projette de transformer le quartier de Hudson Yards en « quartier quantifié ». Grâce à une multitude de capteurs et d'applications proposées aux habitants, l'ensemble de la vie du quartier pourra être quantifiée, analysée et donc optimisée, tel une smart city à échelle réduite. Les appartements pourront être équipés d'éclairages qui s'allument et s'éteignent selon les allées et venues des occupants et l'intensité lumineuse

⁴⁵ NYC Open Data. *Dashboard* [en ligne]. (2017) Disponible sur : < <https://opendata.cityofnewyork.us/dashboard/> > (consulté le 03/06/18)

pourra être ajustée selon l'éclairage naturel de l'appartement. Les immeubles pourront disposer d'ascenseurs dont les moteurs captent l'énergie du freinage pour la réinjecter dans le système électrique du collectif.

Pittsburgh et Chicago améliorer la qualité de vie des citoyens :

La ville de Pittsburgh, en Pennsylvanie, a développé une plateforme open data qui agrège aussi les données des villes environnantes. La ville travaille entre autres sur la création d'une base de données qui détaille au maximum les caractéristiques de chaque parcelle de terrain de la ville et ce en temps réel. L'objectif est de détecter plusieurs choses : fraudes, terrains vacants, risques... Mais aussi d'identifier les causes de la dévalorisation de certains terrains.

A Chicago, l'open data permet entre autres d'améliorer la santé publique. Grâce à « Foodborne Chicago » les tweets et les appels au 311 relatifs à des intoxications alimentaires sont centralisés et permettent aux services d'hygiène d'intervenir avec plus de précision. L'accès aux soins est également facilité par l'intermédiaire d'applications comme « Flu Shots » qui identifie les lieux où la vaccination est gratuite. Enfin, le « Chicago Health Atlas » est un outil potentiel très puissant pour la ville. Cette plateforme centralise des données issues des hôpitaux sur la santé publique et les géo-localise. Ainsi on peut par exemple observer la propagation d'une maladie dans la ville. Prochainement la plateforme pourra proposer une analyse prédictive pour suivre la santé des habitants et identifier de multiples risques. Selon la précision des informations récoltées il sera possible : de suivre la propagation d'une épidémie ou encore d'identifier les secteurs les plus touchés par l'obésité par exemple.

Dernier exemple d'utilisation de la donnée : le dispositif « The Array of Things ». Mené par des architectes et des chercheurs, l'opération consiste à installer des boîtiers dans la ville, qui relèveront plusieurs informations : température, humidité, bruit, pollution, lumière... mais ils permettent aussi d'observer l'activité dans les différents espaces où ils sont implantés, comme par exemple le nombre de piétons grâce au bornage des smartphones. Ces boîtiers pourront permettre d'améliorer la qualité de vie de la population urbaine par l'observation de l'environnement urbain (développement de systèmes d'alertes sur la pollution de l'air, les pollens etc.) ou encore de régler les feux de circulation automatiquement selon le trafic routier et les flux piétons.

Illustration 18 : un capteur du dispositif Array of Things dans les rues de Chicago (source : lafabriquedelacite.com) et installation d'un capteur (source : <https://arrayofthings.github.io>)



3.3. De l'utilité à la nécessité de la gestion du big data : être conscient de l'enjeu pour en éviter les pièges

Bien que l'open data soit une obligation légale en France aujourd'hui, les collectivités reconnaissent également la nécessité de savoir gérer le big data. La multiplication d'applications comme Waze ou encore Airbnb et leur impact sur les territoires illustre bien les problèmes auxquels les collectivités doivent parfois faire face.

Les collectivités dans l'impasse, face à des applications qui impactent directement les territoires - les cas de Waze et Airbnb :

Waze est un GPS collaboratif : les utilisateurs envoient des données en temps réel, qui permettent à l'application de proposer le meilleur itinéraire possible en fonction des flux routiers et des incidents. Les collaborateurs peuvent déclarer un accident, un animal égaré, des routes barrées et même la présence de policiers. Outre la question de signaler les forces de l'ordre sur les routes, l'application cause aujourd'hui un autre problème : en voulant optimiser les trajets, l'application oriente les wazeurs sur de petites routes qui traversent des communes périphériques des grandes villes, engorgeant ainsi leur centre-ville aux heures de pointe. Les voiries deviennent subitement sous dimensionnées, et les municipalités doivent prendre des dispositions pour éviter les embouteillages et les plaintes des riverains qui ne peuvent plus circuler. Mais celles-ci sont souvent démunies, car leur « ennemi » est virtuel (la donnée) et le propriétaire de ces données est Google, une entreprise mondiale contre laquelle il est difficile de lutter. A Lieusaint en Seine et Marne, la commune, envahie par les wazeurs cherchant à éviter les embouteillages, se retrouve obligée d'installer des feux tricolores pour dissuader les automobilistes d'emprunter la route qui traverse la commune⁴⁶.

Aujourd'hui, plusieurs villes sont donc en négociation avec Waze, dont Toulouse. Les municipalités tentent de s'associer avec l'entreprise, afin de mieux gérer la circulation sur leur commune. En octobre 2017 Waze lance une opération de communication avec une proposition aux collectivités : grâce à une application nommée « Connected Citizens » Waze souhaite associer les collectivités et les citoyens dans la récoltes de données relatives au réseau routier. En échange de ces données, Waze s'engage à fournir les données qu'il possède sur l'état du trafic en temps réel. Ces informations sont précieuses pour les villes qui peuvent optimiser leur trafic routier. Dans 70% des cas, les wazeurs signalent un accident de la route avant que l'alerte ne soit donnée aux secours. Il est estimé qu'il serait possible de gagner en

⁴⁶ J.-G.B. avec les éditions départementales. Circulation : Waze, l'appli qui agace les riverains [en ligne]. *LeParisien.fr* (04/06/18) Disponible sur : < <http://www.leparisien.fr/info-paris-ile-de-france-oise/transports/circulation-waze-l-appli-qui-agace-les-riverains-04-06-2018-7752765.php> > (consulté le 05/06/18)

moyenne 4 minutes 30 pour que les secours interviennent⁴⁷, c'est dire l'enjeu que cette application peut représenter pour les pouvoirs publics. En acceptant d'échanger des données et de coopérer avec Waze, les municipalités ont donc la possibilité de réduire considérablement leur trafic routier, mais aussi d'être plus efficace dans les services qu'elles rendent à la population. La coopération entre Waze et la ville de Boston a permis de réduire de 18% la congestion aux plus grands carrefours de la ville et d'ajuster la durée des feux selon le trafic afin d'optimiser la circulation.

Airbnb, plateforme de location de maisons et appartements entre particuliers, a aussi causé des dérèglements dans les villes. L'application a tout d'abord concurrencé fortement les secteurs de l'hôtellerie et du tourisme, en proposant des locations à prix attractifs. Autre problème, certains utilisateurs ont fait de cette activité leur métier : ils investissent dans des logements qu'ils louent exclusivement par le biais de cette plateforme. Ils favorisent donc la location aux touristes et bloquent les potentiels locataires qui souhaitent résider à l'année dans la ville. Autre conséquence à cela, les prix des loyers augmentent dans certaines zones. Certains quartiers se retrouvent donc désertés par les locaux, entraînant à la chaîne des manques à gagner pour des commerces de proximité qui fonctionnent désormais sur un modèle saisonnier et des villes musées. Un cadre commence donc à être exigé par les villes : Amsterdam a négocié avec Airbnb pour que la plateforme empêche les utilisateurs de louer plus de 60 jours par an⁴⁸. En France la loi limite la location à 120 jours par an, ce qui est beaucoup plus souple.

A l'ère du numérique, les collectivités sont donc confrontées à de nombreux problèmes causés par l'utilisation d'applications. Ces phénomènes sont d'autant plus compliqués à maîtriser qu'ils sont difficiles à anticiper et peuvent avoir un impact considérable sur les territoires, avec une rapidité telle que les municipalités n'ont pas le temps de s'adapter, tout

⁴⁷ ARNULF S. Pourquoi Waze échange des données avec les villes [en ligne]. *L'usine-digitale.fr* (02/11/16) Disponible sur : < <https://www.usine-digitale.fr/article/interview-pourquoi-waze-echange-des-donnees-avec-les-villes.N457832> > (consulté le 03/06/18)

⁴⁸ THEPOT M. Quels problèmes pose réellement Airbnb ? [en ligne]. *Latribune.fr* (05/12/16) Disponible sur : < <https://www.latribune.fr/economie/france/quels-problemes-pose-reellement-airbnb-622374.html> > (consulté le 04/06/18)

comme le cadre juridique qui manque souvent pour aider les territoires à solutionner les problèmes. Maîtriser les données qui arrivent de l'extérieur, c'est donc mieux gérer son territoire et anticiper les difficultés, voire tourner ces nouveaux phénomènes sources de déstabilisation à l'avantage des territoires. Car l'association d'entreprises privées comme Waze, à de grandes villes, peut permettre de développer de nouveaux outils très avantageux pour les villes. La donnée est donc devenue un outil stratégique pour les collectivités, qui mesurent aujourd'hui la nécessité de sa gestion et de son utilisation.

Les limites de l'utilisation de la data et de son ouverture :

Parallèlement à la politique open data menée par le gouvernement, la RGPD menée par l'Union Européenne et applicable depuis mai 2018, renforce la protection des données personnelles et responsabilise les acteurs du traitement de ces données. Entre open data et protection des données personnelles, les collectivités doivent se mettre au pas rapidement sur deux sujets assez complexes et parfois antagonistes : entre ouverture et fermeture (ou du moins limitation dans les diffusions, de certains jeux de données pour respecter la vie privée). D'autant plus que ces deux textes n'ont pas été élaborés de manière concomitante⁴⁹. Les collectivités doivent par exemple mettre à disposition les permis de construire délivrés, mais ces derniers contiennent des informations personnelles relatives au dépositaire : nom, adresse... Ces informations personnelles n'étant pas diffusables, les agents devront occulter toutes les mentions personnelles, ce qui va prendre un temps considérable. Finalement, de nombreux documents risquent de ne pas être publiés, alors qu'ils le devraient. Dans tous les cas, la RGPD sera privilégiée à la loi pour une République Numérique (qui prône l'ouverture) car les entorses à cette loi peuvent être sanctionnés. Mais la RGPD facilitera à long terme la politique open data, car elle demande que « seules les données à caractère personnel nécessaires au regard de la finalité spécifique de l'opération⁵⁰ » soient conservées dans les

⁴⁹ ZIGNANI G. Le RGPD et l'open data peuvent-ils faire bon ménage ? [en ligne]. Lagazettedescommunes.com (26/04/18) Disponible sur : < <http://www.lagazettedescommunes.com/557030/le-rgpd-et-lopendata-peuvent-ils-faire-bon-menage/> > (consulté le 04/06/18)

⁵⁰ Auteur associé. RGPD : protéger les données à caractère personnel dès la conception des traitements [en ligne]. Lagazettedescommunes.com (28/02/18) Disponible sur : <

jeux de données, dès leur élaboration. Ainsi, au fil du temps, les données personnelles deviendront quasiment absentes des jeux de données.

Les limites de l'open data commencent néanmoins à se révéler. La création de nouveaux services pourrait concurrencer les services publics : des questions se posent alors sur leur utilité et leur rôle. Quelles conséquences engendre l'ouverture de la donnée et sa réutilisation par d'autres acteurs, privés ? Quel est alors le rôle des services publics ? Deviennent-ils de simples gestionnaires de services externalisés ? Des recompositions seront sans doute à l'œuvre dans les prochaines années. Comme le pointait un rapport de la FING (Fondation Internet Nouvelle Génération) en 2011 : « ces questions ne peuvent pas s'appréhender de manière théorique. Les réponses n'émergeront que de manière progressive, à l'expérience »⁵¹.

La création d'emplois et d'entreprises liées à l'open data a aussi ses limites. Il faut rappeler que les entreprises ou les emplois créés par l'open data et le numérique en général peuvent être précaires, notamment à cause des effets de mode, avec des applications qui peuvent tomber en désuétude, mais également par l'effet de la concurrence. D'autre part, la forme même que prennent ces entreprises (startups, petites entreprises) engendre des contrats précaires comme des contrats à durée déterminée, qui adaptent la main d'œuvre aux besoins des entreprises, trop fragiles et récentes pour proposer des contrats à long terme. Il existe aussi un risque de voir les grandes entreprises étrangères utiliser massivement et abusivement les données ouvertes, sans que cela ne bénéficie à l'Etat et aux collectivités. C'est une économie où tout peut aller très vite, aussi bien positivement que négativement, c'est pourquoi le tissu économique local doit continuer à être diversifié. Cela amène à un questionnement sur le rôle de l'acteur public, comme le souligne A. Courmont : « (...) doit-il

<http://www.lagazettedescommunes.com/551840/rgpd-protéger-les-données-a-caractère-personnel-des-la-conception-des-traitements/> > (consulté le 11/06/18)

⁵¹ FING. *Guide pratique de l'ouverture des données publiques territoriales* [en ligne]. (janvier 2011) Disponible sur : < <http://doc.openfing.org/RDPU/GuidePratiqueDonneesPubliquesv1beta.pdf> > (consulté le 04/05/18)

produire de la donnée uniquement pour ses propres besoins ou prendre en considération les besoins externes d'entreprises privées ? »⁵².

De surcroît, l'ouverture des données peut permettre à des individus mal intentionnés de créer des sites discriminants ou intrusifs. Aux USA, l'ouverture de données ethniques sur la population a donné lieu à la création de cartographies avec des origines ethniques par quartier⁵³. Les types d'applications comme celles développées à Chicago, qui permet de donner une appréciation aux agents de la ville, peuvent aussi basculer dans la dénonciation.

L'analyse prédictive, en plein développement, pose également question. En Afrique du Sud, l'assureur santé Discovery permet de bénéficier de tarifs préférentiels⁵⁴. Ces tarifs sont définis par la récolte de données relatives à la santé de l'assuré par l'intermédiaire d'un bracelet ou bien de ses cartes fidélité au supermarché, qui donnent des informations sur ses achats et donc son alimentation. Ce genre d'utilisation de données à des fins prédictives peut être contesté : un mauvais état de santé peut être causé par une multitude de paramètres, qui n'ont rien à voir avec les données récoltées ou qui ne sont même pas encore connus.

Il convient également de se questionner sur l'utilisation de la donnée. La ville peut-elle être contrôlée par des données, des chiffres qui sont produits et analysés pour guider les habitants et prédire ce qu'il va se passer, tout cela depuis un bureau ? Il semble primordial de rappeler que le big data comporte des données chiffrées, froides. Elles sont certes très utiles pour chiffrer, quantifier, analyser en prenant du recul, ce qui permet parfois d'observer des phénomènes jusqu'alors non identifiés. Mais pour autant, il ne faut pas négliger la nécessité des données sensibles et l'importance de la connaissance du terrain. Car ce sont les êtres

⁵² COURMONT A. Open Data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique [en ligne]. *Informations sociales*, 2015, n°191, pages 40 à 50. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-40.htm> > (consulté le 19/06/18)

⁵³ CHERON A. L'open data : les risques potentiels de l'ouverture des données publiques [en ligne]. *Lejournaldunet.com* (01/03/11) Disponible sur : < <https://www.journaldunet.com/ebusiness/expert/49672/l-open-data---les-risques-potentiels-de-l-ouverture-des-donnees-publiques.shtml> > (consulté le 13/06/18)

⁵⁴ FREZAL S. Outils prédictifs : contresens et dérives [en ligne]. *Esprit* (février 2017) Disponible sur : < <https://esprit.presse.fr/article/frezal-sylvestre/outils-predictifs-contresens-et-derives-39220> > (consulté le 13/06/18)

humains qui font la ville, les personnes, les ménages ne sont pas des numéros. Les citoyens habitent la ville, la pratiquent et l'expérimentent, or cela représente une masse de connaissances que des chiffres ne peuvent pas combler. La donnée produite par les collectivités est dédiée à un usage précis, pour mener une action précise : « toute les données contiennent une certaine représentation du monde social, déterminée par la volonté des acteurs d'intervenir sur celui-ci »⁵⁵.

Lorsque la ville de Boston effectue des croisements entre le nombre de vitres cassées dans la ville et les indicateurs qui renseignent la pauvreté, l'échec scolaire ou encore la violence, cela peut mener à des analyses un peu hâtives. La corrélation de ces chiffres n'est peut-être pas un hasard, mais est-ce suffisant pour prendre le risque de stigmatiser des quartiers ? En Suède, la plateforme Lexbase met en ligne des données sur la criminalité, dont les jugements d'affaires criminelles à l'échelle d'un quartier : la question de la stigmatisation de certains quartiers se pose là encore. Dans ces cas-là, il semble judicieux d'articuler données quantitatives et données qualitatives, avec un travail de terrain, qui permettra d'établir avec certitude les faits relevés par les seuls chiffres.

La question de l'échelle a donc aussi son importance : la donnée diffusée à l'échelle d'un quartier peut provoquer des phénomènes de ségrégation socio-spatiale. Quand un ménage choisit de s'installer dans une ville, le choix de son quartier de résidence peut être réalisé selon des cartes en API, qui leur permettent de visualiser la donnée pour juger du cadre de vie de leur futur lieu de vie. S'ils ont la possibilité de combiner des données « classiques » comme les stations de bus et métro, les commerces de proximité... avec des données de criminalité, la collectivité prend alors le risque de ne pas renouveler la population de certains quartiers et de les voir s'enliser dans la violence et la pauvreté en ne concentrant que des populations qui n'ont pas les moyens de sortir de ces secteurs, voire d'être désertés peu à peu.

⁵⁵ COURMONT A. Open Data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique [en ligne]. *Informations sociales*, 2015, n°191, pages 40 à 50. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-40.htm> > (consulté le 19/06/18)

De surcroît, ces chiffres peuvent être utilisés à des fins politiques. La donnée est donc manipulable à des fins qui ne sont pas toujours dans l'intérêt collectif, c'est pourquoi il faut la manier avec prudence, car bien souvent les chiffres annoncés vont marquer l'esprit de la population.

Enfin, certains chercheurs remettent en cause la notion d'empowerment pour les citoyens⁵⁶. Ce n'est pas parce que les données sont ouvertes qu'elles sont aisément accessibles à tous : certains jeux de données sont complexes (numéro SIREN, bases PERVAL par exemple) et nécessitent un minimum de prérequis pour pouvoir les utiliser. Les collectivités passent d'ailleurs souvent leurs bases de données en revue afin de les simplifier avant de les ouvrir, mais cela a deux inconvénients. Le premier est que cela est chronophage pour les agents, le second est de récupérer une donnée retravaillée qui nuance la réalité, en supprimant des chiffres qui auraient pu nuire à l'image de la collectivité. La sélection des données est donc souvent limitée par un enjeu politique : les données sensibles ne sont pas diffusées. Conséquence relevée par S. Goëta et C. Mabi : « les données publiées en open data parviennent rarement à révéler des fait jusqu'alors inconnus ». Un journaliste du *Monde* leur a d'ailleurs fait part que les données ouvertes sont bien souvent « inoffensives ». En ce sens, le travail des acteurs associatifs et militants est important car il permet de traiter les informations et de les mettre à disposition d'un plus large public.

L'open data risquerait même de donner non pas plus de pouvoir aux citoyens, mais à ceux qui en possèdent déjà : « empower the empowered », c'est la thèse émise par le chercheur canadien M. Gurstein⁵⁷. L'ouverture de la donnée, selon les politiques actuellement menées, tendrait à faciliter le développement des acteurs qui dominent déjà (grandes entreprises par exemple) au détriment des autres.

Enfin, des freins quant à la réutilisation de la donnée peuvent être pointés. Le premier est le manque d'informations complémentaires à un jeu de données. En effet, la donnée est

⁵⁶ GOËTA S., MABI C. L'open data peut-il (encore) servir les citoyens ? *Mouvements* [en ligne]. 2014, n°79, pages 81 à 91. Disponible sur : < [https://www.cairn.info/revue-mouvements-2014\)3\)page-81.htm](https://www.cairn.info/revue-mouvements-2014)3)page-81.htm) > (consulté le 20/06/18)

⁵⁷ GURSTEIN M. *Open data. Empowering the empowered or effective data use for everyone ? First Monday*, 16(2), 2011

produite avant tout pour l'usage de la collectivité, mais elle n'est pas pensée et donc construite pour faciliter une réutilisation pour un autre usage⁵⁸. Un second frein identifié par A. Courmont est la restriction des jeux de données au territoire qui la libère : le réutilisateur est alors limité dans l'usage qu'il peut en faire. D'où l'importance que les services publics travaillent en coopération, afin de mettre en commun leurs données avant de les publier, ce qui permettrait de publier des bases plus riches, complètes et utilisables à une plus grande échelle.

Des propositions pour favoriser l'usage de la donnée par le grand public :

La première proposition, donnée par S. Goëta et C. Mabi (2014), consiste à diffuser la donnée sous deux formats. Le premier serait un format brut, riche en informations, mais qui s'adresse à un public plus averti pour être exploité. Le second serait simplifié, avec une forme allégée et plus accessible.

Une autre proposition de ces deux chercheurs est de renforcer le cadre juridique en matière d'open data, afin d'influencer la sélection de la donnée à ouvrir. Cela dans l'objectif d'ouvrir plus, mais d'inciter également les administrations à diffuser des données qu'elles peuvent juger sensibles. Les auteurs préconisent la création d'un « Conseil National de l'Information Publique » pour arriver à cela et permettre d'ouvrir plus, tout en préservant la vie privée des citoyens.

Enfin, les chercheurs conseillent de mettre l'accent sur la formation et la sensibilisation du public aux enjeux et à l'utilisation de la donnée. Cette donnée, qui envahit le quotidien, doit être comprise de tous. Pour cela, des espaces numériques ouverts au grand public peuvent être mis en place, avec des personnes ressource pour guider les personnes. Dès l'école, l'utilisation de l'open data pourrait être valorisée avec par exemple l'utilisation de données sur l'environnement ou la citoyenneté, afin de faire prendre conscience à chacun de la possibilité qu'offre la donnée si l'on s'y intéresse.

⁵⁸ COURMONT A. Open Data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique [en ligne]. *Informations sociales*, 2015, n°191, pages 40 à 50. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-40.htm> > (consulté le 19/06/18)

Utiliser la donnée c'est aussi apprendre à développer un esprit critique et à chercher par soit même l'information. Cela est d'autant plus important, dans une société où la théorie du complot et le manque de discernement de certains citoyens commence à gagner du terrain.

Pour A. Courmont, la réutilisation des données ouvertes est conditionnée par trois opérations censées renforcer la « solidité » des données, en amont de leur diffusion. Le premier est leur enrichissement⁵⁹ : le producteur de la donnée doit la penser pour qu'elle soit réutilisée, en l'amendant avec d'autres informations. Or aujourd'hui la donnée est pensée et produite pour l'usage unique du service qui la produit. Le second critère est la standardisation. Cela passe d'abord par une homogénéisation des jeux de données entre les services publics en amont de leur diffusion, ce qui revient donc à améliorer la coopération entre les producteurs de données. Cette étape facilite la réutilisation à des échelles plus grandes que celle des territoires producteurs de la data. Enfin, la dernière étape consiste à articuler les jeux de données entre eux, grâce à une donnée pivot, qui permet de les agréger plus facilement et donc de démultiplier les croisements possibles et donc les réutilisations. Cette donnée pivot peut être un code INSEE ou IRIS par exemple. Dans tous les cas, plus elle sera universelle, plus la base de données sera réutilisable pour différents usages. Toutefois il ne faut pas oublier que ces traitements auront un coût financier pour les collectivités.

⁵⁹ COURMONT A. Open Data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique [en ligne]. *Informations sociales*, 2015, n°191, pages 40 à 50. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-40.htm> > (consulté le 19/06/18)

CONCLUSION

La donnée a envahi notre société : elle est partout, dans notre quotidien et influence nos modes de vie. La compréhension du vocabulaire qui l'entoure est parfois difficile, liée à de nombreux anglicismes : data, big data, open data... mais aussi à des termes informatiques, licences, API, formats, plus ou moins obscurs. Combiné à une image de suite de chiffres et d'algorithmes complexes, la donnée peut faire peur et on peut avoir l'impression de se noyer dans une masse d'informations sans rien y comprendre. Pour autant, elle impacte grandement les territoires, que l'on essaie d'y échapper ou non. Le législateur a pris conscience des enjeux liés à la donnée. Il a donc tout d'abord incité les collectivités à ouvrir leurs données, notamment avec la loi NOTRe et la loi pour une République Numérique. Mais il a également tenu à encadrer l'utilisation qui peut en être faite, afin de protéger la vie privée des citoyens.

L'écosystème de la donnée réunit une multitude d'acteurs issus de la sphère publique, associative, privée ou encore civile. Tout le monde est concerné et aujourd'hui tout le monde peut participer à enrichir la donnée et l'utiliser, via des plateformes collaboratives comme Open Street Map, des applications comme Waze ou Citizens Connect... La donnée sert l'action publique en lui apportant des informations essentielles à la gouvernance des territoires : recensement de la population, connaissances topographiques, météorologiques, données économiques, sociales etc. Plus il y a de données à disposition, plus d'informations seront produites et plus de nouvelles connaissances seront produites. Tout cela se démultiplie grâce au partage et à l'ouverture des données.

A Toulouse, l'ouverture de la donnée est lancée au cours de l'année 2011. Un climat favorable prenait place au niveau national et la métropole toulousaine figurait parmi les premières villes françaises à se lancer dans une politique open data. De nombreux échanges, plus ou moins formels, entre les collectivités ont permis de créer un véritable réseau, jusqu'à la création de l'association d'Open Data France. Une multitude de questions se posent lorsque les données s'ouvrent : leur identification, leur structuration, leur format, la licence qui sera utilisée, la manière d'animer le mouvement... Chaque collectivité reste maître de tous ces choix. Malgré plusieurs obstacles à leur ouverture : récupération des jeux de données, mise à

jour, vérification... La donnée, son ouverture et son utilisation représentent un réel atout pour les territoires.

L'open data permet à l'action publique d'être plus efficace, de travailler en transversalité, de gagner du temps mais aussi de l'argent. Le développement d'applications par des acteurs privés aide grandement à réaliser ces économies. L'ouverture des données bénéficie également au tissu économique de la ville, avec la création de nouvelles entreprises qui peuvent faire diminuer le taux de chômage. Des innovations sociales peuvent émerger par l'appropriation du nouveau bien commun que constitue la donnée : participation à des projets urbains, collaboration pour mener des projets d'occupations temporaires d'espaces en friche, pour signaler de désagréments urbains et trouver une solution qui émane de la sphère citoyenne. En outre, elle favorise la création d'un lien de confiance et établi un nouveau dialogue entre citoyen et administration, par la transparence qu'elle procure à l'action publique. De nouveaux services urbains, qui optimisent la ville, peuvent être créés : fluidification de la circulation automobile, économies d'énergies, analyse prédictive... La prochaine étape pour les collectivités françaises sera celle de l'évaluation des politiques publiques engagées, grâce à la donnée récoltée.

Néanmoins, l'utilisation de la donnée a aussi ses limites. Les applications développées, si innovantes soient elles, peuvent causer des dommages aux territoires. C'est le cas de Waze qui engorge les routes secondaires, par report des flux des grands axes vers des cheminements plus confidentiels et donc vite saturés, ou encore d'Airbnb qui concurrence le secteur de l'hôtellerie et perturbe l'accès au logement pour les populations locales. Face à ces changements, entraînés parfois de manière brutale, la ville doit s'adapter et répondre par ses propres moyens, faute d'un cadre législatif trop peu réactif.

D'autres limites peuvent être identifiées : la création d'emplois précaires, la stigmatisation de certains quartiers ou encore les dérives discriminantes de l'utilisation de données en sont des exemples.

La question du rôle du service public va devenir certainement de plus en plus prégnante : si les services sont de plus en plus délégués aux acteurs privés, quelle légitimité restera-t-il à l'acteur public ? Celui d'un simple régulateur ?

Pour conclure, la donnée est un outil précieux : elle permet de chiffrer, quantifier, analyser les phénomènes urbains en prenant du recul. Mais pour fonctionner, les données numériques doivent être couplées des données sensibles et d'expériences d'usage de la ville, afin de certifier les problèmes ou les phénomènes identifiés par le big data. Un territoire ne peut donc pas être administré depuis un bureau, devant un écran d'ordinateur et une multitude de chiffres. Par ailleurs, l'analyse prédictive, qui peut être un outil fascinant, ne doit pas aveugler l'action publique : le futur d'êtres humains, ceux qui fabriquent la ville, ne peut pas être défini par des algorithmes et des probabilités.

La donnée est donc un bien commun qui offre des opportunités qui semblent infinies. Elle est aujourd'hui identifiée comme un outil stratégique pour les collectivités : la gestion et l'utilisation de la donnée est nécessaire et leur assure la maîtrise de leur territoire. Mais les questions éthiques (protection de la vie privée) et démocratiques (accès à la donnée et compréhension) ne doivent pas être négligées, afin que ce terreau de potentielles nouvelles connaissances que constitue la donnée, soit utilisé de manière durable et juste pour tous.

BIBLIOGRAPHIE

ARNULF S. Pourquoi Waze échange des données avec les villes [en ligne]. *L'usine-digitale.fr* (02/11/16) Disponible sur : < <https://www.usine-digitale.fr/article/interview-pourquoi-waze-echange-des-donnees-avec-les-villes.N457832> > (consulté le 03/06/18)

Auteur associé. RGPD : protéger les données à caractère personnel dès la conception des traitements [en ligne]. *Lagazettedescommunes.com* (28/02/18) Disponible sur : < <http://www.lagazettedescommunes.com/551840/rgpd-protoger-les-donnees-a-caractere-personnel-des-la-conception-des-traitements/> > (consulté le 11/06/18)

Better Government Competition. *Apps for Democracy* [en ligne]. (19/07/10) Disponible sur : < <http://bgc.pioneerinstitute.org/apps-for-democracy/> > (consulté le 05/06/18)

CHASSAGNE, C. (dir.). *Open data - ouverture, exploitation, valorisation des données publiques*. Bourgoin Jallieu : territorial éditions, 2017, 134 p.

CHERON A. L'open data : les risques potentiels de l'ouverture des données publiques [en ligne]. *Lejournaldunet.com* (01/03/11) Disponible sur : < <https://www.journaldunet.com/ebusiness/expert/49672/l-open-data---les-risques-potentiels-de-l-ouverture-des-donnees-publiques.shtml> > (consulté le 13/06/18)

CHIGNARD S. *Comprendre l'ouverture des données publiques*. Quercy : éditions FYP, 2012, 191 p.

COURMONT A. Open Data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique [en ligne]. *Informations sociales*, 2015, n°191, pages 40 à 50. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-40.htm> > (consulté le 19/06/18)

Datagalaxy.com. *Cartographie des données personnelles : l'indispensable chantier pour démarrer avec GPDR* [en ligne]. (2018) Disponible sur : < <https://www.datagalaxy.com/blog/gdpr/cartographie-donnees-personnelles-commencer-gdpr/> > (consulté le 14/06/18)

Data.gouv.fr. *Licences de réutilisation* [en ligne]. Disponible sur : < <https://www.data.gouv.fr/fr/licences> > (consulté le 10/05/18)

Economie.gouv.fr. *République numérique : ouverture des données d'intérêt général* [en ligne]. (22/09/2016) Disponible sur : < <https://www.economie.gouv.fr/republique-numerique-ouverture-donnees-d-interet-general> > (consulté le 01/06/18)

Entretien réalisé en juin 2018 avec Sandrine MATHON, chef du service administration à la direction des systèmes d'information de la ville de Toulouse et de Toulouse Métropole, responsable de l'Open Data de la Métropole.

European Commission. *L'open data, c'est quoi ?* [en ligne]. (2018) Disponible sur : < <https://www.europeandataportal.eu/elearning/fr/module1/#/id/co-01> > (consulté le 10/05/18)

FING. *Guide pratique de l'ouverture des données publiques territoriales* [en ligne]. (janvier 2011) Disponible sur : < <http://doc.openfing.org/RDPU/GuidePratiqueDonneesPubliquesv1beta.pdf> > (consulté le 04/05/18)

FREZAL S. Outils prédictifs : contresens et dérives [en ligne]. *Esprit* (février 2017) Disponible sur : < <https://esprit.presse.fr/article/frezal-sylvestre/outils-predictifs-contresens-et-derives-39220> > (consulté le 13/06/18)

GEFFRIER F., MARTINEZ G., LEROY A. *RGPD : que va changer la nouvelle loi sur les données personnelles ?* [en ligne]. (2 mai 2018) Disponible sur : < <http://www.europe1.fr/technologies/rgpd-que-va-changer-la-nouvelle-loi-sur-les-donnees-personnelles-3641006> > (consulté le 04/06/18)

GOËTA S. Un air de famille : les trajectoires parallèles de l'open data et du big data. *Informations sociales* [en ligne]. 2015, n°191, pages 26 à 34. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-26.htm> > (consulté le 20/06/18)

GOËTA S., MABI C. L'open data peut-il (encore) servir les citoyens ? *Mouvements* [en ligne]. 2014, n°79, pages 81 à 91. Disponible sur : < [https://www.cairn.info/revue-mouvements-2014\)3\)page-81.htm](https://www.cairn.info/revue-mouvements-2014)3)page-81.htm) > (consulté le 20/06/18)

Gouvernement.fr. « Pour une République Numérique » [en ligne]. (15/05/17) Disponible sur : <https://www.gouvernement.fr/action/pour-une-republique-numerique> (consulté le 09/05/18)

GUALINO, J. *Dictionnaire pratique informatique, Internet et nouvelles technologies de l'information et de la communication*. Paris : Gualino, 2006, 507 p.

GURSTEIN M. *Open data. Empowering the empowered or effective data use for everyone ? First Monday*, 16(2), 2011

JEGOU, L. Une carte animée du trafic des transports en commun de grandes villes du monde [en ligne]. *Mappemonde*, 2017, n°122. Disponible sur : < <http://mappemonde.mgm.fr/122int1/> > (consulté le 20/06/18)

J.-G.B. avec les éditions départementales. Circulation : Waze, l'appli qui agace les riverains [en ligne]. *LeParisien.fr* (04/06/18) Disponible sur : < <http://www.leparisien.fr/info-paris-ile-de-france-oise/transports/circulation-waze-l-appli-qui-agace-les-riverains-04-06-2018-7752765.php> > (consulté le 05/06/18)

La Fabrique de la Cité. *Comment optimiser les villes avec les data ?* [en ligne]. (23/03/2015) Disponible sur : < [https://www.lafabriquedelacite.com/fabrique-de-la-cite/data.nsf/951F68AEB0B75B1CC1257E0F003225D2/\\$file/etude_def_urbandata2.pdf](https://www.lafabriquedelacite.com/fabrique-de-la-cite/data.nsf/951F68AEB0B75B1CC1257E0F003225D2/$file/etude_def_urbandata2.pdf) > (consulté le 02/06/18)

Larousse. Base de données [en ligne]. *Encyclopédie Larousse*. Disponible sur : <http://www.larousse.fr/encyclopedie/divers/base_de_donnees/185906 > (consulté le 10/05/18)

Larousse. Open data [en ligne]. *Encyclopédie Larousse*. Disponible sur : https://www.larousse.fr/dictionnaires/francais/open_data/10910872 (consulté le 10/05/18)

Laurence. C'est quoi la smart city ? Une introduction compréhensible [en ligne]. *OpenDataSoft* (29 avril 2016) Disponible sur : <<https://www.opendatasoft.fr/2016/04/29/cest-quoi-la-smart-city-une-introduction-a-la-ville-intelligente/> > (consulté le 23/05/18)

Le blog d'Etalab (Etalab). *Etalab* [en ligne]. Disponible sur : <https://www.etalab.gouv.fr/qui-sommes-nous> (consulté le 09/05/18)

LEGIFRANCE. *Projet de loi relatif à la protection des données personnelles* (JUSC1732261L) [en ligne]. (2018) Disponible sur : <https://www.legifrance.gouv.fr/affichLoiPreparation.do;jsessionid=AD5660270AD9F70B94275AC823321680.tplgfr22s_3?idDocument=JORFDOLE000036195293&type=contenu&id=2&typeLoi=proj&legislature=15 > (consulté le 04/06/18)

LeLab, CHRONOS, OUI SHARE. *Quelles coopérations public-privé à l'ère de la data ?* [en ligne]. In : Institut ville durable - Les dessous de l'innovation, n°2. (mai 2017) Disponible sur : < http://institutvilledurable.fr/wp-content/uploads/2017/10/CHRONOS-ENQUETES-n2_VF-VERSION-WEB.pdf > (consulté le 30/05/18)

LiberTIC. *Sensibilisation Open Data*. In : SlideShare [en ligne]. (4 avril 2017) Disponible sur : <<https://fr.slideshare.net/libertic/sensibilisation-open-data> > (consulté le 12/05/18)

MINISTERE DU NUMERIQUE. *LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique*. [en ligne]. Journal officiel, n°0235 du 08 octobre 2016. Disponible sur :

<<https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=3EFDF130F9E755A0B3DFC0AD340AAABB.tplgfr21s3?cidTexte=JORFTEXT000033202746&categorieLien=id> > (consulté le 04/06/18)

NODE-LANGLOIS F., EGLOFF E. Nice inaugure on « boulevard connecté » [en ligne]. LeFigaro.fr (18/06/13) Disponible sur : <<http://www.lefigaro.fr/conjoncture/2013/06/17/20002-20130617ARTFIG00709-nice-inaugure-son-boulevard-connecte.php> > (consulté le 15/06/18)

NYC Open Data. *Dashboard* [en ligne]. (2017) Disponible sur : < <https://opendata.cityofnewyork.us/dashboard/> > (consulté le 03/06/18)

OLLION E. L'abondance et ses revers. Big Data, open data et recherches sur les questions sociales [en ligne]. *Informations sociales*, 2015, n°191, pages 70 à 79. Disponible sur : < <https://www.cairn.info/revue-informations-sociales-2015-5-page-70-.htm> > (consulté le 19/06/18)

OpenData Lab Occitanie. *Lexique autour de l'open data* [en ligne]. (2018) Disponible sur : <<http://www.opendatalab.fr/l-opendata/lexique-autour-de-l-open-data>> (consulté le 24/05/18)

OpenData Lab Occitanie. *L'open data, c'est quoi ?* [en ligne]. (2018) Disponible sur : < <http://www.opendatalab.fr/l-opendata/qu-est-ce-que-l-opendata> > (consulté le 30/03/18)

ORYGA. *L'historique des lois de protection des données en France* [en ligne]. (16 janvier 2018) Disponible sur : < <https://www.oryga.com/actualites-rgpd/historique-lois-protection-des-donnees-france/> > (consulté le 04/06/18)

PRIOL J. *Le big data des territoires – Open data, protection des données, smart city, civic tech, services public... Les nouvelles stratégies de la donnée au service de l'intérêt général*. Buijsière : éditions FYP, 2017, 223 p.

REPUBLIQUE FRANCAISE. *LOI n° 2015-991 du 7 août 2015 portant nouvelle organisation territoriale de la République* [en ligne] Journal officiel n°0182 du 8 août 2015 page 13705. Disponible sur : < <https://www.legifrance.gouv.fr/eli/loi/2015/8/7/2015-991/jo/texte> > (consulté le 04/06/18)

REPUBLIQUE FRANCAISE. *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés* [en ligne] Journal officiel du 7 janvier 1978 page 227. Disponible sur : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460&categorieLien=id> (consulté le 04/06/18)

REPUBLIQUE FRANCAISE. *LOI n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal* [en ligne] Journal officiel, du 18 juillet 1978 page 2851. Disponible sur : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000339241&categorieLien=id> (consulté le 04/06/18)

Secrétariat d'État à la réforme de l'Etat et à la simplification. *Pour une action publique transparente et collaborative : plan d'action national pour la France 2018-2020* [en ligne]. (2017) Disponible sur : < <https://www.etalab.gouv.fr/wp-content/uploads/2018/04/PlanOGP-FR-2018-2020-VF-FR.pdf> > (consulté le 03/05/18)

Secrétariat général pour la modernisation de l'action publique. *Rapport au premier ministre sur la gouvernance de la donnée 2015 – la donnée au service de la transformation de l'action publique* [en ligne]. (décembre 2015) Disponible sur : < https://www.gouvernement.fr/sites/default/files/document/document/2016/01/rapport_au_premier_ministre_sur_la_gouvernance_de_la_donnee_2015.pdf > (consulté le 01/06/18)

THEPOT M. Quels problèmes pose réellement Airbnb ? [en ligne]. *Latribune.fr* (05/12/16) Disponible sur : < <https://www.latribune.fr/economie/france/quels-problemes-pose-reellement-airbnb-622374.html> > (consulté le 04/06/18)

THEVENIN L. *Groupama exploite l'open data pour la sécurité routière* [en ligne]. *Les Echos.fr* (20/04/17) Disponible sur : < https://www.lesechos.fr/20/04/2017/LesEchos/22429-132-ECH_groupama-exploite-l-open-data-pour-la-securite-routiere.htm > (consulté le 23/06/18)

Thibaud. L'Energy Atlas d'Amsterdam : vers la ville durable grâce à l'urban data [en ligne]. *OpenDataSoft* (22/02/17) Disponible sur : < <https://www.opendatasoft.fr/2017/02/22/lenergy-atlas-damsterdam-vers-la-ville-durable-grace-a-lurban-data/> > (consulté le 06/06/18)

Toulouse Métropole. *Concours défi numérique : défiez les données ouvertes par Toulouse Métropole* [en ligne]. (2014) Disponible sur : < https://www.toulouse-metropole.fr/projets/les-projets/-/asset_publisher/VCxskVEr4zgQ/content/concours-defi-numerique-defiez-les-donnees-ouvertes-par-toulouse-metropole;jsessionid=E6B4D84B77E04F58E202FAA9CD4777BC > (consulté le 01/06/18)

Toulouse Métropole. *#CP – Des élèves d'EPITECH Toulouse remportent le hackathon smart city de Toulouse* [en ligne]. (18/12/15) Disponible sur : < <https://www.toulouse-metropole.fr/-/-cp-test-des-eleves-de-epitech-toulouse-remportent-le-hackathon-smart-city-de-toulouse> > (consulté le 28/05/18)

ZIGNANI G. Le RGPD et l'open data peuvent-ils faire bon ménage ? [en ligne]. *Lagazettedescommunes.com* (26/04/18) Disponible sur : < <http://www.lagazettedescommunes.com/557030/le-rgpd-et-lopendata-peuvent-ils-faire-bon-menage/> > (consulté le 04/06/18)



La donnée est partout : capteurs sur les voiries, smartphones, objets connectés... elle a envahi notre quotidien. Acteurs privés, publics, société civile : tout le monde produit de la donnée, parfois sans même s'en rendre compte. Cette donnée est aujourd'hui identifiée comme un outil derrière lequel se cachent de nombreux bénéfices à en tirer : développement économique, innovation sociale, gain de temps, d'efficacité...

Pour l'action publique, l'enjeu consiste aujourd'hui à ouvrir cette donnée, afin qu'elle soit réutilisée. Les raisons de cette ouverture sont multiples : souci de transparence et de renouveau dans le dialogue entre citoyen et administration, mais aussi de développement économique et innovation.

Depuis peu, l'open data (ouverture de la donnée) est devenu obligatoire pour les collectivités de plus de 3 500 habitants. Le législateur tente d'accompagner l'ouverture de la donnée dans les territoires, grâce à la mission Etalab créée en 2011. Mais il doit aussi gérer la protection de ces données et notamment celles concernant la vie privée, d'autant plus que cet aspect a été renforcé avec le Règlement Général sur la Protection des Données impulsé par l'Union Européenne en mai 2018.

A travers l'exemple de la métropole toulousaine, ainsi que d'autres expériences étrangères, ce mémoire propose de comprendre en quoi la donnée, et son ouverture, constitue un outil stratégique pour les collectivités. Car la maîtrise de la donnée est aujourd'hui une nécessité pour les territoires, afin de mieux les gérer. Toutefois, l'ouverture des données comporte aussi des risques et des limites, qu'il convient d'identifier afin d'en éviter les pièges.

Mots clé : Open data – Collectivités – Donnée – Service public – Territoires