



HAL
open science

Livrable 1.2 Analyse de portails métropolitains de données ouvertes à l'échelle internationale

Françoise Paquienséguy, Valentyna Dymytrova

► To cite this version:

Françoise Paquienséguy, Valentyna Dymytrova. Livrable 1.2 Analyse de portails métropolitains de données ouvertes à l'échelle internationale . [Rapport de recherche] 1.2, Equipe d'accueil lyonnaise en Sciences de l'information et de la communication. 2017. hal-01449348

HAL Id: hal-01449348

<https://hal.science/hal-01449348>

Submitted on 30 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



elico

Équipe de recherche de Lyon en sciences
de l'information et de la communication

**Analyse de portails métropolitains
de données ouvertes
à l'échelle internationale**

LIVRABLE n°2

Rendu le 27 juin 2016 et finalisé le 30 janvier 2017

Version janvier 2017

Direction scientifique : F. Paquienséguy (PU- 71ème)

Version	Date	Contributeurs
Mars 2016	25/03/2016	Françoise Paquienséguy (PU) Valérie Larroche (MCF) Marie-France Peyrelong (MCF) Valentyna Dymytrova (Ingénieur de recherche)
Juin 2016	26/06/2016	Françoise Paquienséguy (PU) Valérie Larroche (MCF) Marie-France Peyrelong (MCF) Valentyna Dymytrova (Postdoc)
Janvier 2017	30/01/2017	Françoise Paquienséguy (PU) Valentyna Dymytrova (Postdoc)

INTRODUCTION

Porté par le laboratoire Elico (EA 4147), ce livrable n°2 est associé à la tâche 2.1 « *Definition and understanding of the existing tools & needs in a smart city* ». Il rend compte de l'analyse de plusieurs portails métropolitains de données ouvertes, menée de février à décembre 2016. L'objectif de ce travail a été d'identifier les différents modèles de portails Open data à l'échelle internationale afin de mieux comprendre la place des données liées à la mobilité au sein d'un large panel de portails. Dans cet objectif, nous avons constitué un corpus de 24 portails Open data métropolitains, 12 français et 12 étrangers. Le choix des portails a été opéré en fonction de la taille des métropoles et de leurs avancées en matière de l'ouverture des données.

L'analyse à la fois sémiologique et stratégique a été menée en trois temps. D'abord, nous avons étudié chaque portail en fonction d'une grille d'analyse établie qui accorde une attention particulière aux stratégies de communication des portails (Annexe 1). Ensuite, nous avons comparé les portails entre eux afin de faire ressortir des grandes tendances dans les modalités de diffusion des données ouvertes. Enfin, nous avons analysé les thématiques regroupant les jeux de données de la mobilité et les applications liées à celle-ci afin de comprendre la place réservée à ce type de données par les portails et leurs concepteurs.

Les premiers résultats de l'analyse de l'équipe Elico ont été restitués le 27 juin 2016 à la Métropole de Lyon, puis le 12 janvier à l'ANR lors de la revue de mi-parcours. Ils ont suscité intérêt et questions de la part de nos interlocuteurs. Ils ont été affinés et enrichis jusqu'en janvier 2017 grâce à un travail de veille conduit par la post-doctorante du projet, V. Dymytra.

1. CONTEXTE ET ÉTAT DE L'ART

Les portails Open data sont des matérialisations des démarches d'ouverture des données publiques, fortement impulsées par plusieurs directives européennes (directive PSI, 2003 et directive Inspire, 2007) et cadres législatifs nationaux récents (loi Macron, 2015, loi NOTRe, 2015, loi Valter, 2015 et loi Lemaire, 2016). Même si les collectivités les plus importantes en nombre d'habitants et capacités budgétaires sont les plus avancées dans l'ouverture des données, celle-ci s'impose aujourd'hui à toutes les communautés de plus de 3500 habitants.

Les portails des données ouvertes sont sensés remplir plusieurs objectifs liés à la modernisation de l'action publique : celui de transparence dans la relation des collectivités avec leurs citoyens, celui de potentiel économique et d'innovation pour les entreprises locales et nationales et celui de participation et d'implication des citoyens dans la vie des collectivités. Ces portails constituent une interface entre les collectivités, les différents fournisseurs et les usagers des

données. Cette dernière catégorie est loin d'être uniforme et il convient de distinguer les usagers intermédiaires ou professionnels comme, par exemple, développeurs, data journalistes, data scientifiques ou administrations elles-mêmes et les usagers finaux ou amateurs qui sont les citoyens. Les différents usagers des données ouvertes se distinguent en fonction des objectifs qu'ils se fixent mais aussi par rapport à leurs compétences en matière de la manipulation des données.

2. METHODOLOGIE

2.1. CONSTITUTION DU CORPUS

Débutées dans le monde anglo-saxon au travers d'initiatives telles que les portails *data.gov* et *data.gov.uk* aux Etats-Unis et en Angleterre, les démarches d'ouverture des données sont profondément inscrites dans les contextes nationaux, à la fois juridiques, politiques, organisationnels et techniques. L'étude comparée des portails métropolitains d'Open data à l'échelle nationale, européenne et internationale permet de saisir les variations et les invariants dans les modalités de diffusion des données ouvertes.

Les portails d'Open data ont été sélectionnés par rapport à la taille des collectivités territoriales et à leurs avancées en matière de l'ouverture des données¹. Au total, pendant plus d'un an, nous avons analysé 24 portails : 12 portails nationaux et 12 portails internationaux, dont les noms et les adresses URL sont listés ci-dessous.

- Montpellier : Open Data Montpellier numérique : <http://opendata.montpelliernumerique.fr/>
- Nantes : Nantes ouverture des données : <http://data.nantes.fr/>
- Nice : Open Data Nice Côte d'Azur : <http://opendata.nicecotedazur.org>
- Rennes : Open Data Rennes : <https://data.rennesmetropole.fr/>
- Paris : Open Data Paris : <https://opendata.paris.fr/>
- Toulouse : Open Data Toulouse Métropole : <https://data.toulouse-metropole.fr/>
- Agen : Open Data Agen : <https://data.agen.fr/>
- Bordeaux : Open Data Bordeaux : <http://opendata.bordeaux.fr/>
- Corse : Open Data Corsica : www.opendata.corsica/
- Île-de-France : Open Data Île-de-France : <http://data.iledefrance.fr>
- Grand Lyon : Data Grand Lyon – La Métropole de Lyon : <https://data.grandlyon.com/>
- Strasbourg : Open data - Données | Je suis un professionnel | Strasbourg.eu : <http://www.strasbourg.eu/ma-situation/professionnel/open-data/donnees>

¹ Fondation Open Knowledge International liste les communes à l'échelle internationale disposant d'un portail des données ouvertes: <http://dataportals.org/>.

- Londres : London Datastore – Greater London Authority: <https://data.london.gov.uk/>
- Amsterdam : Open Data Amsterdam - Gemeente Amsterdam: <https://data.amsterdam.nl/>
- Berlin: Open Data Berlin : <https://daten.berlin.de/>
- Séoul: 서울 열린 데이터 광장 : <http://data.seoul.go.kr/>
- Chicago: City of Chicago | Data Portal: <https://data.cityofchicago.org/>
- Montréal: Portail données ouvertes Montréal: <http://donnees.ville.montreal.qc.ca/>
- New York: NYC Open Data : <https://nycopendata.socrata.com/>
- Helsinki : Helsinki Region Infoshare | Open regional data: <http://www.hri.fi/>
- Milan: [City of Milan Open Data - - Data Portals](#)
- Barcelone: Open Data de l'Ajuntament de Barcelona : <http://opendata.bcn.cat/opendata/>
- Turin : [AperTO - Gli OpenData della Città di Torino](#) : <http://aperto.comune.torino.it/>
- Birmingham : [Birmingham Data Factory - Birmingham City Council](#): <https://data.birmingham.gov.uk/>

2.2 CATÉGORIES D'ANALYSE DES PORTAILS

La grille d'identification des portails présentée en Annexe 1 a pour objectif d'identifier et de fédérer les informations nécessaires à l'analyse sémiologique et stratégique qui en découle. Chaque catégorie a une valeur spécifique pour l'analyse menée.

- Le nom du portail garantit sa visibilité, sa crédibilité et sa notoriété. Son analyse permet de mieux comprendre la vision que la métropole porte sur ses données et leur diffusion. L'adresse URL renvoie à la stratégie du référencement du site, car la présence dans l'URL du mot clé lié à une requête moteur est pour les moteurs de recherche un signe de pertinence de la page.

- La description de la commune permet de mieux comprendre le lien entre la mise en œuvre de l'ouverture de données et la taille de la collectivité en termes de superficie, de caractéristiques sociodémographiques et d'économie.

- Le nom de la plateforme permet d'identifier les plateformes les plus utilisées et de comprendre le lien entre le choix de la plateforme et les fonctionnalités proposées par le portail.

- La date de la création du portail situe la réalisation du site web dans le temps et par rapport à sa propre stratégie de l'ouverture de données.

- Le nombre de jeux de données disponibles permet de mieux évaluer l'avancement de l'ouverture de données.

- Les noms des rubriques renseignent sur les thématiques de données ouvertes et sur les façons de classer les données au sein des catégories.

- Le type de métadonnées indique la qualité de la donnée et la possibilité de sa réutilisation.

- Les formats proposés renseignent sur les types de ré-utilisateur visés (hommes ou machines) et le degré d'interopérabilité de données.

- Le type de licence proposé permet de connaître les conditions d'accès et de réutilisation de données et caractérise la stratégie de l'ouverture de données.

- Les modes d'accès aux données (téléchargement, navigation, API) renseignent sur les réutilisations envisagées par le diffuseur de données et le type de public visé.

- Les fonctionnalités principales du portail concernent l'organisation des contenus et des services offerts et permettent d'identifier les publics visés.

- Le public cible renvoie aux éléments qui orientent le portail vers un type d'utilisateurs précis.

Par ailleurs, nous nous sommes intéressées à la visibilité des données liées à la mobilité sur les portails analysés en intégrant dans la grille deux catégories suivantes : noms des thématiques dans lesquelles trouve des jeux de données de la mobilité et, le cas échéant, noms des applications liées à la mobilité publiées sur le portail. Ces catégories permettent de mieux comprendre la façon dont sont définies les données de la mobilité mais aussi la place que la collectivité réserve à des services créés à partir des données ouvertes publiés.

2.3. ANALYSE

Après avoir constitué le corpus des portails à étudier et avoir établi la grille d'analyse, nous avons procédé à des analyses approfondies de chaque portail. Il est important de prendre en compte le fait que les analyses ont été menées sur un laps de temps assez important, allant du mars au décembre 2016. Entre temps, certains portails d'Open data de notre corpus ont été complètement refaits ou ont disparu. Par exemple, la nouvelle version du portail d'Open data de Rennes lancée en septembre 2016 propose une nouvelle interface et de nouvelles fonctionnalités, très orientées vers la visualisation des données². Quant au portail d'Open data de Montpellier, intitulé *Open Data Montpellier Territoire numérique*, il est devenu le portail *M3M Montpellier Méditerranée Métropole*. Par ailleurs, certaines communes qui publiaient leurs données *via* la plateforme nationale data.gouv.fr se sont dotées de leurs propres portails. Par exemple, Brocas, commune de 700 habitants dans le sud-ouest de la France, a ouvert ses données depuis 2011 *via* la plateforme

²http://metropole.rennes.fr/fileadmin/rrm/documents/Espace_Presse/Communique_de_presse/2016_CP_nouveau_portail_sept.pdf

data.gouv.fr, elle possède aujourd'hui son propre portail, intitulé *Brocas Open Data – L'Open Data*, c'est aussi pour les petites collectivités³.

3. RESULTATS

3.1. PORTAIL VS PLATEFORME

Le portail d'Open data est toujours le front office d'une plateforme, qui assure l'hébergement, la gestion des jeux de données et de leur interopérabilité. Cette distinction, proposée par Philippe Beaune de l'EMSE, a été prise en compte dans l'analyse menée en amenant un questionnaire sur l'impact que le choix de la plateforme pourrait avoir sur les fonctionnalités et les contenus éditoriaux du portail⁴.

L'identification de la plateforme utilisée n'est pas toujours facile, beaucoup de communes font le choix de ne pas l'indiquer sur le portail d'une manière suffisamment explicite. Parmi les portails analysés, plusieurs utilisent les différentes plateformes commerciales (Socrata, OpenDataSoft, Ckan, DataPress) et d'autres sont basés sur les plateformes indépendantes, élaborées par la commune elle-même (Grand Lyon⁵). Certaines villes font le choix de publier les données sur les pages dédiées du site web municipal (Strasbourg). Les petites communes intègrent directement leurs données à la plateforme data.gouv.fr (Brocas).

D'une manière générale, nous avons constaté que la plupart des portails français analysés recourent à la plateforme OpenDataSoft⁶. Grâce à l'interopérabilité de cette plateforme avec la plateforme data.gouv.fr, les données publiées par l'intermédiaire d'un portail OpenDataSoft peuvent être automatiquement référencées sur le portail national et inversement, la plateforme OpenDataSoft peut récupérer des données au niveau national et les publier au niveau local. Pour

³ <http://opendata.brocaslesforges.fr/>.

⁴ Larroche, V., Peyrelong, M-F., Beaune, Ph. « Ouverture des données métropolitaines : quels projets quels enjeux ? », communication, Forum national des Interconnectés, 12 et 13 décembre 2016, Lyon.

⁵ Conçu par le service Innovation Numérique & des Systèmes d'Information de la métropole de Lyon, la plateforme technique est entièrement basée sur des logiciels libres : Linux, Apache/Tomcat, PostgreSQL/PostGIS, MapServer 6. Le catalogue des données est basé sur l'outil Géosource.

⁶ Créée en 2011, OpenDataSoft développe une plateforme en ligne complète conçue pour la transformation rapide et économique de tous types de données en services innovants (APIs, visualisations de données, monitoring en temps réel) sans aucune limite de volumétrie. Sa mission : faciliter la publication, le partage et la réutilisation de données et d'APIs par les utilisateurs métiers. Basée à Paris, OpenDataSoft compte parmi ses clients Total, ERDF, GrDF, Suez Environnement, le Ministère de l'Education nationale, de l'Enseignement Supérieur et de la Recherche, la Région Île-de-France, la Ville et le Comté de Durham (NC, USA), SNCF, Veolia, m2ocity, ou encore Citilog.

encourager les communes hésitantes dans leur démarche, en 2015, OpenDataSoft a généré automatiquement des portails de données pour toutes les villes françaises de plus de 20 000 habitants. En 2016, après l'adoption de la loi NOTRe, la société a automatiquement conçu 5000 portails des données pour les collectivités locales de plus de 3500 habitants. Un site web, www.opendataterritoires.fr, est dédié à cette démarche.

À l'échelle d'autres villes européennes, nous avons constaté l'usage d'autres plateformes, comme par exemple, DataPress pour Londres et Amsterdam et Ckan pour Berlin, Turin et Birmingham, ou encore Socrata pour les portails de New York et de Chicago.

3.2. FONCTIONNALITÉ DES PORTAILS

Quelle que soit la plateforme utilisée, les portails comportent généralement des fonctionnalités similaires.

Les jeux de données sont le plus souvent accessibles à partir d'un moteur de recherche ou d'un classement thématique, souvent matérialisé par des icônes cliquables. Certains portails proposent aussi les nuages des tags qui donnent un aperçu des données disponibles et invitent à leur découverte. Les catalogues des jeux de données proposent la recherche par format de données, fournisseur ou producteur, API, mots-clés, type de licence ou visualisation. Les jeux de données peuvent être le plus souvent triés par pertinence, fréquence de mise à jour ou encore popularité. Par exemple, les données du portail de Chicago peuvent être classées par nombre de vues, nombre de téléchargement ou par notes attribuées par les usagers. Certaines collectivités proposent aussi les statistiques concernant la disponibilité et les usages des données du portail. Par exemple, Nantes et Milan publient les listes des données statiques et des données en temps réel les plus téléchargées.

3.2.1. Présentation de la démarche de l'ouverture des données

Chaque portail est présenté en tant que « concrétisation », au sens de Simondon de l'engagement des collectivités d'ouvrir leurs données. Incontournable, la rubrique « Démarche » (« About », « Project » ou « City Data Strategy ») comprend le plus souvent la définition de l'Open data et présente l'intérêt de la démarche. L'ouverture des données est le plus souvent décrite en tant qu'une nouvelle relation entre les collectivités et les citoyens et une nouvelle opportunité économique pour les entreprises du territoire. Les mots clés mis en avant sur les portails dans les discours d'accompagnement de la démarche sont : la transparence, la participation, le partage, la collaboration, l'innovation, l'accessibilité, l'interopérabilité, l'efficacité, la réutilisation, le développement du tissu économique local et une meilleure qualité de vie pour les habitants. Toutefois, les discours d'accompagnement et de promotion de la démarche diffèrent entre les portails analysés. Certains portails se limitent à la présentation de la vision métropolitaine de la

donnée et de sa mise à disposition (Grand Lyon, Paris). D'autres inscrivent leur démarche métropolitaine dans le mouvement national de l'Open data en se référant au cadre législatif national (Nantes) ou en soulignant leur engagement dans l'association Open Data France (Toulouse, Nice). Rares sont les communes qui présentent les objectifs de mise à disposition des données à la fois par rapport à plusieurs échelles : métropolitaine, régionale, nationale, européenne, voire internationale. Nous pouvons citer à titre d'exemple le portail de Milan qui décrit au sein de la rubrique dédiée le cadre européen, le cadre national (Italie) et le cadre régional (Lombardie) avant de présenter la démarche de l'ouverture des données de la commune. Un autre exemple est le portail des données ouvertes de la Corse dont la réalisation s'inscrit dans le projet européen HOMER visant à développer une communauté d'acteurs de l'Open data sur le bassin méditerranéen en associant 8 pays.

Les formats qu'empruntent les discours de présentation de la démarche sont principalement textuels. Certains portails mettent à disposition des documents-guides téléchargeables au format PDF qui présentent les stratégies de l'ouverture des données ouvertes des métropoles correspondantes (Montréal), des manuels des standards techniques utilisés (Chicago) ou des guides techniques (« Le guide de demande de données publiques auprès des collectivités » et « Le manuel d'utilisation des données de l'Agglo3D » de Bordeaux Métropole).

Toutefois, plusieurs portails français se démarquent par le recours aux vidéos, souvent considérés comme étant des formats plus ludiques. Par exemple, les portails de Nantes et de Nice diffusent le spot pédagogique de présentation de la démarche d'ouverture des données numériques publiques, initié par LiberTIC, soutenu par Nantes Métropole et réalisé par A2B Production en licence Art Libre en 2011. En plus, le portail de Nantes propose une vidéo institutionnelle « L'open data à la nantaise » qui expose la démarche de l'ouverture des données par la voix du président de Nantes Métropole, de la représentante de l'association LiberTIC et la voix *off* qui crée l'effet d'un reportage journalistique. Le portail de Bordeaux comporte une rubrique spéciale « Les vidéos » avec deux documents, « L'open data mode d'emploi » et « Des données et des hommes ». Si la première vidéo explique la navigation du portail en invitant les usagers à manipuler les outils mis à disposition, la deuxième vidéo est un reportage métier, qui raconte l'histoire des données et des hommes en retraçant tout l'itinéraire d'une donnée, depuis sa collecte jusqu'à sa publication en Open data à travers l'exemple des métiers des Parcs & Jardins. La visée pédagogique du choix des vidéos est confirmée avec plusieurs activités autour de la vidéo Pars & Jardins qui permettent d'approfondir le sujet (« Les métiers », « Le saviez-vous ? ») et de vérifier sa bonne compréhension (« Le quiz »).

D'une manière générale, chaque portail présente les acteurs-partenaires du projet de l'ouverture des données. Cela peut se manifester sous une forme iconique compacte (présence des logos) ou sous forme d'une présentation textuelle plus détaillée au sein des rubriques « Acteurs », « Partenaires » ou « Fournisseurs ».

3.2.2. Souci d'actualisation des contenus

La plupart des portails comportent la rubrique dédiée aux actualités. Nommée « Actualité », « Les actus », « News » ou « Highlights », celle-ci comporte des informations sur les évolutions du portail (jeux de données récents, nouvelles réutilisations) et l'avancement du projet de l'ouverture des données (nouveaux partenariats, concours, appels à projets, événements). Par ailleurs, plusieurs portails intègrent un espace dédié aux flux des tweets qui présentent les contenus en lien avec l'ouverture des données publiques. Il s'agit de contenus actualisés en temps réel, mais qui ont une durée de visibilité très réduite, car un nouveau tweet chasse les précédents. Une autre modalité d'actualisation des contenus est la rubrique « Blog » présente dans plusieurs portails analysés (Chicago, Londres). Inscrite dans un dispositif de communication plus informel, celle-ci raconte comment les données ont été récoltées et réutilisées dans de différents contextes et s'ouvre vers les expériences de l'ouverture des données d'autres territoires.

3.2.3. Place et définition de l'utilisateur

Afin de faciliter la navigation sur le site, certains portails proposent des rubriques telles que « Glossaire » et « FAQ ». Les espaces directement dédiés à l'interaction avec les usagers comprennent les rubriques « Contact » et « Forum ». Par ailleurs, plusieurs portails proposent des espaces personnalisés, qui nécessitent l'inscription et l'identification et qui permettent une navigation élargie des données et des outils mis à disposition pour les réutiliser comme par exemple, la création de visualisations ou l'utilisation d'API.

Les usagers des portails peuvent, le plus souvent, commenter et évaluer les jeux de données publiés, ils peuvent aussi parfois discuter des jeux de données sur les forums dédiés. Certains sites intègrent un formulaire *via* lequel les usagers peuvent faire une demande ou une suggestion d'un jeu de données (Nice, Montréal, Berlin ou Amsterdam). Rares sont cependant les portails qui rendent accessibles les demandes formulées. Dans notre corpus de 24 portails, c'est le cas de Toulouse et de Montréal. D'autres portails font le choix des forums de discussion où les usagers peuvent partager leurs attentes et idées et contribuer ainsi à la construction du projet de l'Open data (Rennes). Certains portails invitent les usagers à donner des idées des applications, contribuer au recensement des applications existantes et signaler des erreurs (Barcelone, Turin).

D'une manière générale, la plupart des portails proposent des espaces dédiés aux réseaux socio-numériques qui permettent aux usagers de partager les données, les visualisations et les applications avec leurs contacts par mail ou *via* Facebook, Twitter, G+, Tumblr, Instagram, Pinterest, Flickr, YouTube ou LinkedIn. Certains portails offrent de suivre les actualités du portail en s'inscrivant aux flux RSS ou, plus rarement, à la lettre d'information.

3.2.4. Souci de réutilisation

Tous les portails comportent une rubrique bien visible concernant les licences et les conditions d'utilisations des données publiées. La plupart des portails distinguent les réutilisations professionnelles et les réutilisations amateurs. Cela se manifeste le plus souvent par différentes entrées dans le portail. Par exemple, « Coin des débutants » et « Coin de développeurs » sur le portail des données ouvertes de Grand Lyon. La plupart des portails spécifient des entrées spécialement dédiées aux développeurs par des noms « API » ou « Développeurs ». Rares sont toutefois les portails qui mettent à disposition à la page d'accueil des outils qui permettent aux usagers lambda de créer des cartes (Paris, Toulouse, Rennes, Agen) ou générer des visualisations graphiques (Rennes, Barcelone). Certains portails comportent des rubriques dédiées aux réutilisations qui recensent les applications créées à partir des jeux de données disponibles. Nous y reviendrons plus loin en présentant les applications liées à la mobilité.

2.4. PRÉSENTATION DES JEUX DE DONNÉES

Le nombre de jeux de données disponibles varie en fonction des portails et change constamment. Au moment de notre première observation (mars 2016) et pour notre corpus restreint aux 24 portails, le nombre de jeux de données publiés varie de 27 pour la commune de Brocas (770 habitants) à 180 jeux de données pour Paris (2 229 621 habitants) et 1317 pour la ville de New York (8 550 405 habitants). En janvier 2017, lors de la rédaction du livrable, le nombre de jeux de données passe à 33 pour Brocas, 204 pour Paris et 1656 pour New York.

Le nombre de jeux de données disponibles est un indice important de l'avancement du projet de l'ouverture des données. Pour cette raison, il est presque systématiquement bien visible et bien mis en valeur par la couleur, la typographie ou l'emplacement sur le portail.

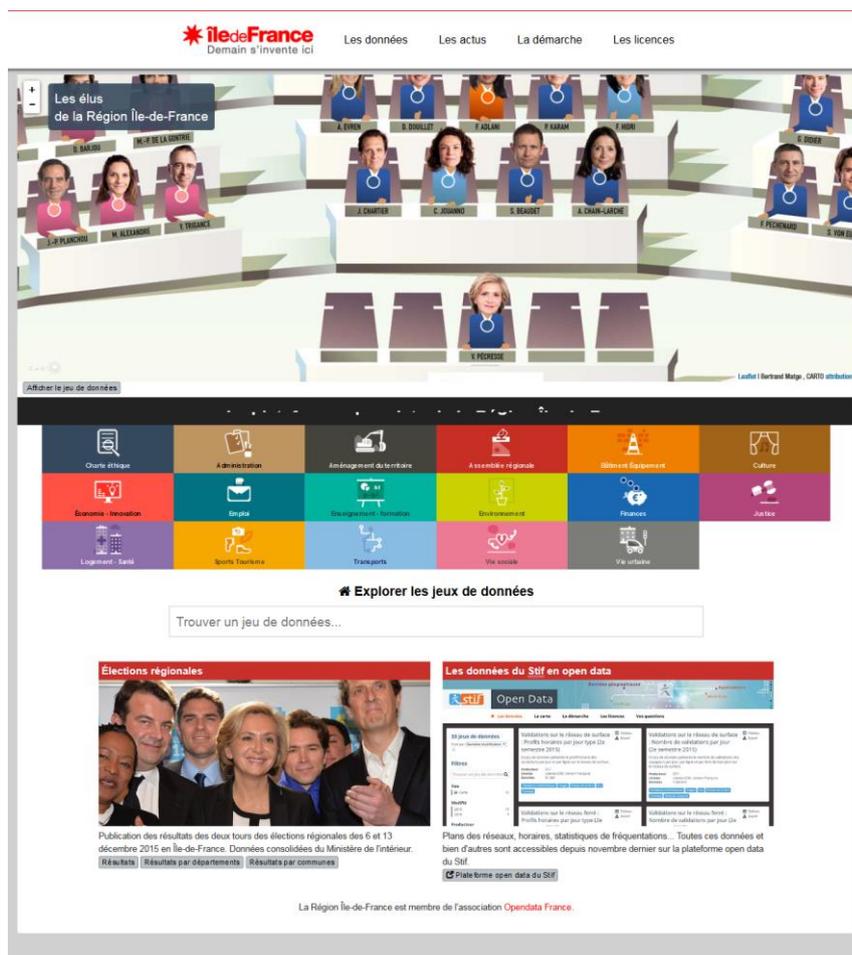
2.4.1. Thématiques des jeux de données

Les données publiées sur le portail sont regroupées au sein des catégories thématiques. Malgré les variétés des dénominations dues aux langues et aux cultures nationales, les thèmes des données sont assez proches. Les thématiques les plus fréquentes sont « Administration », « Culture », « Education », « Environnement », « Economie », « Santé », « Sécurité », « Transport » et « Urbanisme ». Parmi les thématiques plus rares, citons « Elections », « Elimination des déchets », « Charte éthique », « Commerces », « Jeunesse », « Justice », « Lieux saints », « Météo », « Logement et immobilier », « Numérique », « Patrimoine », « Proximité », « Protection des consommateurs », « Solidarité » et « Vie associative ».

2.4.2. Icônes et visualisations

Les thématiques sont souvent accessibles dès la page d'accueil sous forme d'icônes, plus ou moins conventionnelles. Cette façon d'entrer dans les données correspond à la visée d'accessibilité, largement affirmée dans les discours d'accompagnement et de promotion des portails des données ouvertes. En même temps, à partir de la page d'accueil, certains portails représentent les jeux de données sous forme des visualisations graphiques.

Figure 1. Capture d'écran de la page d'accueil du portail d'Île de France (20 janvier 2017).



Par exemple, la page d'accueil du portail des données ouvertes d'Île-de-France (Figure 1, ci-dessus) comporte à la fois une visualisation interactive présentant les élus de la Région en fonction de leur appartenance politique et un aperçu des jeux de données disponibles sous forme des icônes cliquables. Cela rend la page d'accueil attractive mais surtout permet de découvrir une donnée particulière et de montrer sa réutilisation possible.

Toutefois, certains portails préfèrent une sobriété dans la présentation des données à la page d'accueil qui se manifeste par l'absence d'icônes et de visualisations (Paris, Turin, Amsterdam). Ce

choix peut s'expliquer par une orientation vers les usagers avisés, ayant une idée assez précise de ce qu'ils cherchent et peuvent trouver sur le portail.

Figure 2. Capture d'écran du portail des données ouvertes de la Mairie de Paris.

PARISDATA MAIRIE DE PARIS

Connexion

Les données L'API La licence La démarche Cartographe

Bienvenue sur Paris Data
Le site de la démarche Open Data de la Ville de Paris. Vous trouverez ici l'ensemble des jeux de données publiés par les services de la Ville et ses partenaires sous licence ODDL.

Trouver...

Actus: le Partenariat pour un gouvernement ouvert se tient à Paris
Du 7 au 9 décembre 2016, la France accueillera à Paris le 4e Sommet mondial de l'Open Government Partnership ou Partenariat pour un Gouvernement Ouvert. Le Partenariat pour un Gouvernement Ouvert s'attache, au niveau international, à promouvoir la transparence de l'action publique, l'ouverture à de nouvelles formes de consultation et de participation citoyenne et l'utilisation des nouvelles technologies et du numérique pour renforcer la démocratie.

DataCity
INVENONS LA VILLE DE DEMAIN
DataCity est un programme d'open innovation co-conçu par Numa et la Ville de Paris, qui réunit acteurs industriels, startups, et autorités locales pour relever ensemble des challenges urbains afin de rendre la ville plus intelligente, et plus agréable à vivre pour ses citoyens grâce aux données. DataCity s'articule autour de plusieurs étapes, pour une durée totale de 10 mois. Découvrez et rejoignez le programme.

PARIS OPEN INNOVATION
Paris Open Innovation est le groupe Meetup de la démarche Open Innovation de la Ville de Paris. Rejoignez-nous pour être tenu informé des prochaines rencontres et participez à la création d'une ville plus intelligente et durable.

Tweets à propos de @opendataparis

Florian Lainez @overflorian
On attends L'API ouverte avec impatience @EcoEmballagesSA1 @opendataParis @OSM_FR

Tiens @EcoEmballagesSA1 dit vrai, j'ai trouvé des capteurs dans l'engin! @opendataParis @OSM_FR

Thèmes utilisés

- Equipements, Services, Social (Utilisé par 47 jeux de données)
- Mobilité et Espace Public (Utilisé par 39 jeux de données)
- Administration et Finances Publiques (Utilisé par 30 jeux de données)
- Culture (Utilisé par 24 jeux de données)
- Citoyenneté (Utilisé par 22 jeux de données)

Jeux de données les plus populaires

- Stations Vélib - Disponibilités en temps réel (1210281 téléchargements)
- Stations et espaces AutoLib de la métropole pari... (839632 téléchargements)
- Liste des prénoms 2004 à 2015 (19229 téléchargements)
- Les arbres remarquables (11936 téléchargements)
- Liste des sites des hotspots Paris WIFI (11312 téléchargements)

Dernières réutilisations

- Carte des lieux de mémoire Plaques commémoratives 1939-1945 (Il y a 14 jours)
- Analyse temps réel et animation JCDecaux Bike Stations Data (Il y a un mois)
- Toilettes publiques gratuites Sanisettes - Données géographiques (Il y a un mois)
- App iOS de sectorisation scolaire Secteurs scolaires (Il y a 2 mois)
- Carte du risque routier à Paris Accidentologie (Il y a 2 mois)

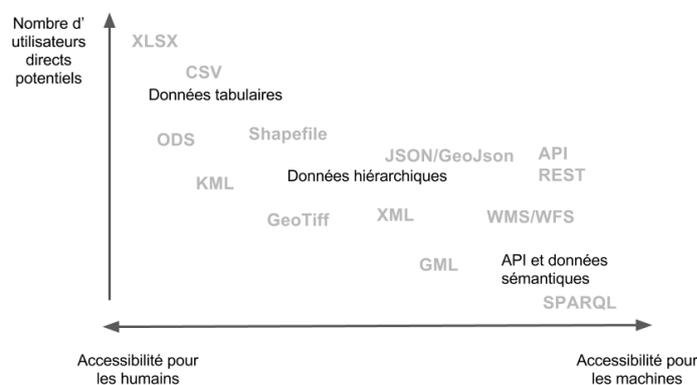
2.4.3. Formats des données

Les formats utilisés pour publier les données ouvertes sont une composante importante de l'accessibilité de l'information et de sa réutilisation. Les portails analysés présentent les données dans une variété de formats.

Tout d'abord, par rapport au degré d'ouverture des formats⁷, nous avons relevé aussi bien l'utilisation de formats ouverts comme le HTML, JSON ou CSV que celles des formats propriétaires comme *Shapefile*, développé par ESRI et le PDF, développé par Adobe.

Ensuite, en fonction des usagers visés (Figure 3, ci-dessous), les portails publient les données tabulaires (CSV, XLSX) ou hiérarchiques (JSON, XML) qui sont destinées aux usagers lambda et les interfaces de programmation (API) qui sont destinés aux développeurs. En effet, 13 parmi 24 portails analysés proposent une entrée spéciale dédiée aux développeurs qui est intitulée « Développeurs » ou « API ». Les fichiers plats les plus répandus sont CSV, JSON, Excel. Quant aux fichiers géographiques, ce sont GeoJSON, Shapefile, KML. Ce point fera l'objet d'un approfondissement *via* l'enquête développeurs conduite par Elico (janvier/avril 2017).

Figure 3. Accessibilité des formats des données pour les humains et pour les machines⁸.



⁷ À l'instar des données, les formats peuvent être différenciés en fonction de leur degré d'ouverture. Ils sont qualifiés d'ouverts lorsque la spécification technique qui permet de lire et de créer ce format est disponible publiquement et gratuitement et lorsque le format est soumis à un mécanisme de maintenance public et ouvert à la participation, propriétaire et fermés. Les formats propriétaires sont ceux dont la spécification technique est disponible publiquement mais dont la maintenance et l'évolution sont assurées par un propriétaire unique. Enfin, les formats sont fermés lorsque la spécification technique n'est pas disponible.

⁸ Graphique issu du portail de données ouvertes de la ville de Montréal. <http://donnees.ville.montreal.qc.ca/portail/formats-de-donnees/>. Consulté le 22 décembre 2016.

Les données sont le plus souvent disponibles en téléchargement, en API ou en consultation sous forme de tableaux, cartes ou de visualisation infographique.

Quant aux types de données, 16 portails sur 24 présentent aussi bien des données statiques que des données dynamiques. Les données dynamiques n'étant pas systématiquement des données fluides ou en temps réel.

2.4.4. Métadonnées

Les fichiers de données publiés sur les portails sont systématiquement accompagnés de métadonnées, qui décrivent la donnée et assurent un accès à celle-ci *via* un répertoire. Leur choix est particulièrement important pour la compréhension des données et leur interopérabilité.

La plupart des métadonnées des portails analysés comportent les catégories suivantes :

- le nom du fichier, parfois accompagné de quelques lignes de description précisant comment la donnée a été collectée ou préconisant les utilisations possibles ;
- le thème ;
- la liste des mots-clés ;
- les tags ;
- le (-s) nom (-s) du producteur/propriétaire/gestionnaire/diffuseur de la donnée ;
- la date de la publication initiale et la date de la dernière mise à jour ;
- la fréquence de mise à jour ;
- les formats disponibles ;
- la licence ;
- la langue ;
- la couverture géographique.

D'une manière moins systématique, les métadonnées peuvent comporter le modèle de données, les réutilisations, les statistiques relatives à la donnée (nombre de réutilisations ; nombre d'abonnés ; note attribuée), les contacts du propriétaire des données, la source, le lien hypertextuel avec la référence ou la couverture temporelle.

2.4.5. Types de licences

Pour encourager la réutilisation des données, tous les portails précisent clairement sous quelle (s) licence(s) les données publiées sont utilisables. Les portails analysés recourent principalement à

deux types de licences : licences dérivées de l'ODbl (Open Database Licence)⁹ de l'Open Knowledge Foundation et celles dérivées de CC BY¹⁰ de Creative Commons.

La licence ODbl repose sur des principes simples d'utilisation et de partage à l'identique pour les données et les bases de données. Elle autorise le partage, la création et l'adaptation à condition de mentionner la paternité, de faire le partage en conditions identiques et de faire perdurer le principe d'Open data.

Les licences *Creative Commons* sont fondées sur le droit d'auteur. La licence CC BY autorise toute exploitation, y compris à des fins commerciales, ainsi que les créations dérivées, dont la distribution sont également autorisées sans restriction, à condition de l'attribuer à son auteur en citant son nom.

Parmi 12 portails français analysés, seul celui de Nice utilise la « Licence Ouverte / Open Licence »¹¹ conçue par Etalab dans le cadre de la politique du Gouvernement en faveur de l'ouverture des données publiques. Cette licence facilite et encourage la réutilisation des données publiques mises à disposition gratuitement et reste compatible avec les standards des licences Open Data développées à l'étranger, notamment celles du gouvernement britannique (*Open Government Licence*) ainsi que les autres standards internationaux (ODC-BY, CC-BY 2.0).

Par ailleurs, un seul portail dans notre corpus, celui de Grand Lyon propose ses propres modèles de licences : licence ouverte, licence engagée et licence associée, définies en fonction de réutilisation des données (Figure 4, ci-dessous).

Figure 4. Types de licences du portail des données ouvertes du Grand Lyon¹².

	 LICENCE OUVERTE	 LICENCE ENGAGÉE	 LICENCE ASSOCIÉE
Utilisation commerciale de données			
Mention de la source et de la date de mise à jour			
Déclaration sur l'usage de la donnée			
Authentification nécessaire avant téléchargement			
Possibilité d'une redevance			

⁹ <http://opendatacommons.org/licenses/odbl/>. Consulté le 23 décembre 2016.

¹⁰ <https://creativecommons.org>. Consulté le 23 décembre 2016.

¹¹ <https://www.etalab.gouv.fr/licence-ouverte-open-licence>. Consulté le 23 décembre 2016.

¹² <https://data.grandlyon.com/connaître-nos-licences/>. Consulté le 23 décembre 2016.

2.4. PUBLIC CIBLE

Les discours de présentation et de promotion des portails des données ouvertes affirment que ceux-ci s'adressent à tout le monde, aussi bien aux hommes qu'aux machines, pour reprendre le slogan du portail de la ville de Berlin : « *Offene Daten lesbar für Mensch und Maschine. Das ist das Ziel* »¹³. Les publics le plus souvent mentionnés sur les portails analysés sont : citoyens, habitants, entreprises, étudiants, chercheurs, développeurs, journalistes, associations et même, d'autres acteurs publics.

Figure 5. Différenciations des publics cibles des portails.

Indicateurs	Public amateur	Public développeur
Différentes entrées	Onglet « Coin de débutants » ou absence d'onglet particulier	Onglets : « API », « Développeurs », « Coin de développeurs » Espace d'identification
Présentation des données dès la page d'accueil	Icones thématiques Exemples de visualisations	–
Disponibilité des outils facilitant la réutilisation	Outils de création de cartes et de graphiques	API
Choix des formats des données	Formats tabulaires ou hiérarchiques	API
Éléments pédagogiques	Manuels d'utilisation Glossaire/Dictionnaire de données ouvertes FAQ Vidéos	-
Applications	Disponibilité des applications Formulaire de suggestion de nouvelles applications	-

Les analyses des fonctionnalités et des interfaces des portails des données ouvertes permettent de différencier les éléments qui concernent davantage le public amateur sans compétences techniques particulières et le public professionnel ayant des compétences d'analyse et de réutilisation des données. Nous avons présenté les indicateurs les plus importants témoignant de telles orientations du portail dans le tableau ci-dessus (Figure 5).

¹³¹³ En français : « Les données ouvertes lisibles par les hommes et les machines. C'est l'objectif ».

3. LE CAS DES DONNÉES LIÉES À LA MOBILITÉ

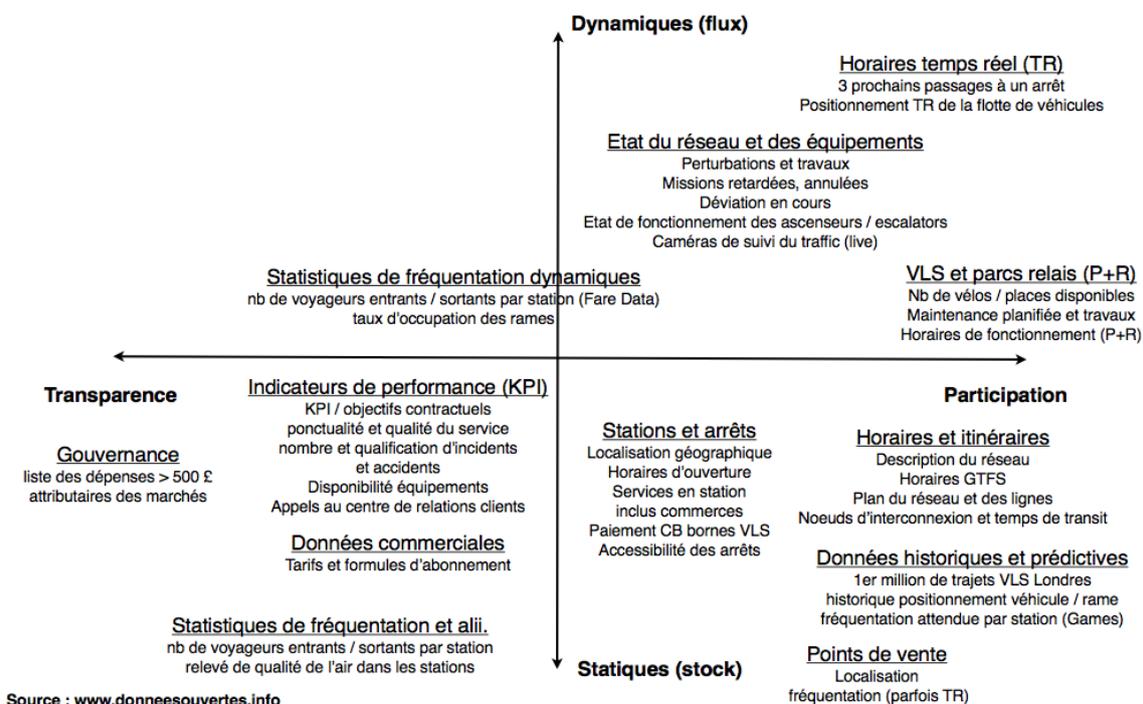
Encouragée par le cadre juridique récent (Loi Macron, 2015 et Loi Lemaire, 2016), l'ouverture des données liées à la mobilité est considérée comme étant stratégiquement importante pour les usagers. Les portails analysés comportent tous des données liées aux transports et mobilité. Nous avons aussi bien les différents types de données (descriptives, géographiques, historiques, prédictives, statistiques, en temps réel) que les données concernant différents modes de déplacement (voiture, transports publics, vélo, marche), le stationnement (parkings) et les conditions de circulation (enneigement, verglas, travaux).

3.1. VISIBILITÉ DES DONNÉES DE LA MOBILITÉ

En raison de leur nature très variée (Figure 6, ci-dessous), les données de la mobilité se distribuent sur le portail entre les thématiques diverses.

Figure 6. Les données ouvertes dans le domaine des transports publics¹⁴.

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



¹⁴ <https://donneesouvertes.info/category/articles/mobilite-transport/>. Consulté le 2 janvier 2017.

Nous retrouvons le plus souvent les données de la mobilité au sein des catégories : « Transports », « Mobilité », « Trafic », « Stationnement », « Déplacement » et parfois aussi « Espace public-voierie », « Accessibilité », « Equipement », « Aménagement » et « Urbanisme ».

3.2. APPLICATIONS LIÉES À LA MOBILITÉ

Plusieurs portails analysés accordent une attention particulière aux applications, présentées comme des exemples concrets de la réutilisation des données ouvertes. En même temps, les portails ne portent pas de traces ou d'exemples de réutilisation des données par des data journalistes, data scientifiques et citoyens lambda.

14 parmi 24 portails analysés publient la liste des applications mobiles développées à partir des données ouvertes et proposant aux citoyens des services permettant d'améliorer leur quotidien. Toutefois, lorsqu'on s'intéresse aux applications liées à la mobilité, nous remarquons que celles-ci ne sont pas présentes sur tous les 14 portails identifiés. En effet, seulement dix portails comportent plusieurs applications de ce type : Nantes, Rennes, Nice, Bordeaux, Toulouse, Montréal, Helsinki, Amsterdam, Berlin et New York. Les applications référencées sur les portails sont de plusieurs types : Android, Apple, Web, Dataviz ou Windows. Elles peuvent être gratuites ou payantes. Par exemple, le portail des données ouvertes de la ville de Nantes publient les applications gratuites, comme Parkopedia à côté des applications payantes, comme Parking Nantes 44. Certains portails se limitent à la liste des applications, d'autres en précisent les fonctionnalités, donnent des informations sur les développeurs/propriétaires et sur les conditions d'utilisation ou encore affichent les notes et les commentaires attribués par les usagers.

La plupart des applications référencées sur les portails se limitent au territoire en question (commune, métropole ou région) et sont développés par des start-ups, des entreprises privées ou des développeurs indépendants lors des *hackatons* organisés par ces collectivités territoriales. Directement accessibles aux citoyens, les applications servent d'intermédiaire dans l'accès aux informations publiées sur les portails. Les informations-services que les citoyens recevront *via* les applications résultent des traitements, indexations et interprétations des éditeurs de ces applications, lesquels visent une rentabilité économique. Les données ouvertes, souvent présentées dans les formats techniques nécessitant les compétences particulières ne paraissent pas directement destinées à des citoyens.

CONCLUSION

L'analyse des portails métropolitains des données ouvertes à l'échelle internationale révèle une réelle diversité des de types, de formats, de métadonnées et de thématiques des jeux de données disponibles. Les discours d'accompagnement de la démarche de l'ouverture des données et les choix ergonomiques et visuels des portails varient également. Au-delà des spécificités des contextes nationaux, à la fois juridiques, politiques, organisationnels et techniques, les portails des données ouvertes expriment le plus souvent une orientation envers des publics cibles précis, professionnels ou amateurs. Ces derniers sont rarement envisagés en tant que ré-utilisateurs potentiels de l'Open data mais plutôt en tant qu'usagers des services créés par des intermédiaires professionnels à partir des données disponibles.

Belles vitrines de la modernisation de l'action publique et de la promotion du territoire, les portails ne rendent pas compte des difficultés administratives, juridiques, techniques, économiques et politiques auxquelles l'ouverture des données est confrontée et des stratégies complexes des différents acteurs concernés par ce processus.

BIBLIOGRAPHIE DU LIVRABLE

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

Flichy, P. (2013), « Rendre visible l'information. Une analyse sociotechnique du traitement des données », *Réseaux*, vol. 2, n° 178-179, p. 55-89.

Kitchin, R. (2014), *The Data Revolution: Big data, Open data, data infrastructures and their consequences*, London, Sage.

Labelle, S., Le Corf, J-B. (2012), "Modalités de diffusion et processus documentaires, conditions du « détachement » des informations publiques. Analyse des discours législatifs et des portails open data territoriaux". *Les Enjeux de l'Information et de la Communication*, n° 13/2, 2012, p. 59-71.

Larroche, V., Peyrelong, M-F., Beaune, Ph. « Ouverture des données métropolitaines : quels projets quels enjeux ? », [Forum national des Interconnectés](#), 12 et 13 décembre 2016, Lyon.

Paquienséguy, F. (2016a), « A qui profite le chiffre? Les portails métropolitains Open Data ». In: Avenati, O., Chardel, P-A. (dir.) *Datalogie. Formes et imaginaires du numérique*, Editions Loco, p. 126-135.

Paquienséguy, F. (dir.) (2016b), *Open data : Accès, territoires, citoyenneté : des problématiques info-communicationnelles*, Paris, Editions des Archives contemporaines.

Thorsby, J., Stowers, N.L. G., Wolslegel, K., Tumbuan, E. (2016), Understanding the content and features of open data portals in American cities, *Government Information Quarterly*. En ligne: <http://www.sciencedirect.com/science/article/pii/S0740624X16301071>.

Zuiderwijk, A. et al. (2012), "Socio-technical Impediments of Open Data", *Electronic Journal of E-Government*, vol. 10, n° 2. En ligne: <http://bit.ly/17yf4pM>.

ANNEXE 1. GRILLE DE RECUEIL DES DONNÉES DES PORTAILS

Date de réalisation de la fiche :

Description de la commune ou du propriétaire de la plateforme :

Adresse pour le portail :

Concepteurs de la plateforme :

Date de conception de la plateforme :

Fonctionnalités :

Nombre de jeux de données global :

Nombre de jeux de données par rubrique (en moyenne) :

Noms des rubriques :

Métadonnées :

Données statiques / dynamiques :

Formats proposés :

Type de licences :

Modes d'accès (téléchargement, navigation, API) (par rubrique?) :

Nombre de jeux sur la mobilité :

Noms des thématiques où on trouve des jeux sur la mobilité :

Public cible de la plateforme :

Entrée réutilisateurs / développeurs ou pas (Existence d'un profil qui donne une entrée directe re-utilisateur) :

Applications ou pas (si oui lesquelles) :

Jeux de données parking ? (si oui lesquels ?) :