



# **GUIDE PRATIQUE DE L'OUVERTURE DES DONNEES PUBLIQUES TERRITORIALES**

## **Guide à l'usage des territoires et de leurs partenaires**

***VERSION DE TRAVAIL N° 1***  
***Janvier 2011***

**Note :** ce guide pratique se fixe pour objectif de faciliter l'engagement des collectivités territoriales dans une démarche d'ouverture de leurs données publiques. Il se fonde sur une observation des expériences françaises et européennes, et notamment sur les démarches des Communautés urbaines de Bordeaux et Rennes. Cette version doit être considérée comme une version de travail, encore imparfaite. Elle sera enrichie et améliorée tout au long de l'année 2011.

Nous accueillons avec plaisir vos commentaires, remarques et contributions par email ou bien en ligne à l'adresse : <http://fing.org/vademecum2011>

Amandine BRUGIÈRE, [abrugiere@fing.org](mailto:abrugiere@fing.org)  
Charles NÉPOTE, [charles.nepote@fing.org](mailto:charles.nepote@fing.org)

# Table des matières

|  |           |
|--|-----------|
| <b>Introduction - Contexte .....</b>   | <b>3</b>  |
| Qu'est-ce que l'ouverture des données publiques? .....   | 3         |
| Pourquoi ce guide pratique ?.....  | 4         |
| Pourquoi y a-t-il un processus d'ouverture des données publiques aujourd'hui ? .....                             | 5         |
| <b>1. Ouvrir et utiliser des données publiques territoriales.....</b>  | <b>10</b> |
| 1.1. Une donnée publique, c'est quoi au juste ?.....   | 10        |
| 1.2. Qu'attend-on de l'ouverture et la réutilisation des données publiques ? .....                               | 12        |
| 1.3. Un nouvel écosystème d'acteurs interdépendants.....   | 27        |
| 1.4. Un cadre juridique très favorable .....   | 34        |
| <b>2. Comment s'y prendre ? .....</b>  | <b>36</b> |
| 2.1. Engager une politique de réutilisation des données publiques.....   | 36        |
| 2.2. Définir sa politique d'ouverture et de diffusion des informations publiques : les grands choix à faire..... | 39        |
| 2.3. Comment rendre les données accessibles ? .....  | 47        |
| <b>Annexes .....</b>   | <b>55</b> |
| A1. "L'ouverture des données et après..." .....  | 55        |
| A2. Petit lexique de l'OpenData .....  | 61        |
| A3. Répertoire des acteurs, organismes et institutions.....  | 63        |
| A4. Bibliographie - Webographie .....  | 66        |

# INTRODUCTION – CONTEXTE

## QU'EST-CE QUE L'OUVERTURE DES DONNEES PUBLIQUES?

Partout en Europe, les Etats et les collectivités locales s'engagent dans l' "ouverture" de leurs données publiques. Le sujet s'invite dans les campagnes électorales nationales, dans les programmes stratégiques des Ministères<sup>1</sup> comme des collectivités territoriales, et se diffuse de plus en plus communément dans la presse.



Mais de quoi s'agit-il au juste ?

- De mettre à disposition des citoyens, des entreprises, des chercheurs, des associations et même, des autres acteurs publics, les masses de données numériques que les collectivités produisent dans leur activité quotidienne : informations géographiques, statistiques, localisation des services, textes réglementaires, études, mesures, barèmes, informations transports, annuaires, etc.
- Et d'en permettre la réutilisation par d'autres logiciels, applications et services.

Pour quoi faire ? Améliorer les services aux habitants ou en inventer de nouveaux, faire émerger de nouvelles connaissances, enrichir le débat public...

<sup>1</sup> Cf. le volet numérique du Plan de relance France numérique 2012 <http://www.francenumerique2012.fr/>

## **POURQUOI CE GUIDE PRATIQUE ?**

---

L'ouverture des données publiques n'est pas une option pour les acteurs publics : elle est rendue obligatoire par des directives européennes et une législation française qui en a fait un droit opposable. Citoyens, associations, entreprises, sont en droit d'exiger que les "données publiques" leur soient délivrées pour qu'ils en fassent leur propre usage, y compris commercial.

Mais cette ouverture constitue également une chance. En s'y engageant, les acteurs publics ont l'occasion de gagner en efficacité, en mutualisant leurs propres bases de données : combien de départements d'une même administration, combien d'administrations d'un même territoire, dupliquent-ils les mêmes bases de données, les mêmes cartes, les mêmes statistiques, évidemment pas tout à fait cohérentes entre elles ?

L'ouverture des données publiques peut également contribuer au développement économique et à l'amélioration de nombreux services aux habitants comme aux entreprises : en permettant à des acteurs de proximité de mieux répondre aux besoins particuliers de tel quartier, de telle catégorie de population, de tel bassin d'emploi, on crée de l'activité tout en améliorant la qualité de vie.

L'accès aux données publiques offre enfin l'opportunité de faire émerger des connaissances inédites sur un territoire, et de partager les éléments de la décision publique avec une société civile qui le demande – et qui, sinon, se fera sa propre idée en dehors de tout dialogue avec les acteurs publics.

Il s'agit donc d'une véritable opportunité pour les territoires, qui plus est moins coûteuse et moins complexe qu'il n'y paraît de prime abord. Chaque territoire la saisira à sa manière, en expérimentant et généralisant. Malgré tout, de nombreux éléments doivent être pris en compte : juridiques, techniques, économiques, organisationnels. Plusieurs territoires en Europe, et en France, montrent la voie : leur expérience a vocation à servir aux autres.

Ce Guide pratique s'inscrit dans la démarche d'accompagnement des territoires initiée par la FING autour de la Réutilisation des données publiques<sup>2</sup>. Ce guide se donne pour objectifs d'apporter des éclairages et d'inviter les acteurs publics à agir. Des moyens, des outils sont déjà à portée de main, des initiatives existantes, en France comme à l'étranger, constituent déjà des bonnes pratiques, ou à tout le moins des sources d'enseignement.

Il s'enrichira, toute l'année 2011, des retours d'expériences et d'initiatives des différents territoires engagés.

C'est le moment de se lancer : "Si vous ne le faites pas vous-mêmes, d'autres le feront à votre place !".

---

<sup>2</sup>

<http://fing.org/?-Reutilisation-des-donnees,138->

## **POURQUOI Y A-T-IL UN PROCESSUS D'OUVERTURE DES DONNEES PUBLIQUES AUJOURD'HUI ?**

---

L'ouverture et la réutilisation des données publiques s'inscrit dans une continuité historique :

- L'informatisation des administrations depuis plusieurs décennies : bureautique, dématérialisation des documents administratifs et des procédures, constitution de base de données et de systèmes d'information géographiques, publication sur le web, téléservices numériques, etc. ;
- Des politiques publiques successives de modernisation de la relation aux citoyens, de transparence de l'information publique, d'accès aux documents, etc.
- Enfin un mouvement plus large d'évolution du web et de l'interrelation des informations qui s'y trouvent publiées.

### **D'une législation basée sur l'accès aux documents administratifs à une législation basée sur la réutilisation**

Dès 1978, l'accès aux documents administratifs a été incité et encadré par une loi (Loi n° 78-753 du 17 juillet 1978), afin d'instaurer une relation de transparence entre les administrations et les citoyens, et d'améliorer l'accès aux droits. Tout document détenu par l'administration, qu'il soit ou non produit par elle peut être demandé par un citoyen, directement ou (en cas de refus) par l'intermédiaire de la [CADA](#) (Commission d'accès aux documents administratifs).

Cette loi a été modifiée par l'ordonnance 2005-650 du 6 juin 2005 qui transpose en droit français la directive européenne de 2003 portant sur la réutilisation des informations du secteur public.

Ainsi la préoccupation initiale centrée sur la transparence et l'accès aux droits s'étend à un objectif d'innovation économique et sociale.

Du point de vue de la Commission européenne, il s'agit d'inviter les administrations européennes à soutenir plus fortement le développement économique, et à créer une synergie entre secteur public et secteur privé sur le marché de l'information<sup>3</sup>.

Mais l'objectif de la Commission, comme des acteurs publics qui s'engagent concrètement dans cette démarche, est plus large. Les données, statiques ou dynamiques forment la matière première de la production de connaissances sur le territoire et ceux qui y vivent, de la décision et de la production de services. Dans nos sociétés complexes, ceux-ci s'appuient sur un nombre croissant d'informations issues de sources diverses. Echanger des données devient indispensable. Les acteurs publics ont besoin d'informations produites par d'autres administrations, par les entreprises, les chercheurs ou les associations – et vice versa. Ainsi, en rendant leurs données réutilisables, les collectivités franchissent un pas supplémentaire, à la fois dans la recherche de leur propre efficacité, et dans leur relation avec le reste de la société.

A côté du droit démocratique à la transparence que couvre l'accès aux documents publics, émerge donc un droit économique et social de *réutilisation* des données : l'administration considère comme son devoir de distribuer la capacité de faire des usages pertinents de ses propres données.

Cependant, cette évolution pose à son tour de nouvelles questions auxquelles il faut apporter des réponses pratiques : qui peut réutiliser quelles données, et sous quelles contraintes ? Que faire si la mise à disposition à un coût ? Dans quelle mesure la responsabilité de l'acteur public se trouve-t-elle engagée par les usages qui sont faits de ses données ? Comment faire en sorte que l'ouverture des données, dont on ne peut évidemment pas contrôler tous les usages, produise des résultats globalement positifs d'un point de vue social, économique, environnemental et démocratique ?

---

<sup>3</sup> Ce mouvement est initié dès 1989 avec la publication du document "Lignes directrices pour améliorer la synergie entre secteur public et secteur privé sur le marché de l'information".  
[http://ec.europa.eu/information\\_society/policy/psi/index\\_en.htm](http://ec.europa.eu/information_society/policy/psi/index_en.htm)

## EN FRANCE, UNE TRADITION ANCIENNE DE PUBLICATION DES DONNEES JURIDIQUES

Il est souvent de bon ton d'affirmer que la France est en retard sur ses voisins européens en matière d'ouverture de ses données. Certes, d'autres pays tels que le Royaume-Uni apparaissent en pointe, mais la France a su se montrer pionnière dans le domaine, et le mouvement y est aujourd'hui bien engagé.

Depuis les années 80, la France est engagée dans la mise à disposition de ses données juridiques, afin de faciliter l'accessibilité et l'intelligibilité de la loi. Si "nul n'est censé ignoré la loi", ce devoir citoyen s'avère une tâche impossible au regard des 8 000 lois et plus de 110 000 décrets en vigueur. Les pouvoirs publics français ont pour cette raison mis en place dès 1984 un "service public de diffusion du droit", constituant des bases de données officielles pour les informations juridiques (lois, règlements, jurisprudence), et accessibles aux citoyens<sup>4</sup>. Les avancées technologiques ont été largement utilisées par les producteurs et éditeurs des bases de données juridiques officielles. En 2002 un décret<sup>5</sup> instaurant la mise en ligne gratuite sur l'internet des informations juridiques marque un nouveau tournant. (Notons que le gouvernement britannique vient d'ouvrir en octobre 2010, son site d'informations juridiques "[Acts and statutory instruments](#)<sup>6</sup>").

L'information publique fait l'objet d'une politique volontariste sur l'internet depuis déjà plus de 10, avec le portail [service-public.fr](#), créé en 2000 (et son guide des droits et démarches) ; ainsi que de [vie-publique.fr](#), le site d'information et de documentation des grands débats publics.

En revanche, cette diffusion porte aujourd'hui sur des documents formatés, éditorialisés et publiés sur des pages web. La mise à disposition de données "brutes" (budgétaires, géographiques, statistiques, de service...), directement issues des systèmes informatiques des administrations, représente une nouvelle étape que les acteurs publics s'apprêtent aujourd'hui à franchir.

## Les "open data", un mouvement mondial

Le mouvement d'ouverture des données ("Open Data") ne se limite pas aux données publiques. Un grand nombre de secteur d'activités, ayant besoin de récolter et d'analyser des masses de données (la recherche, la statistique, le secteur de l'environnement, le journalisme, la culture, l'intelligence économique, la finance, l'urbanisme, la cartographie, pour ne citer qu'eux) ont développé depuis longtemps des pratiques de partage de données, selon des modalités propres.

**Le monde scientifique** a, par exemple, largement contribué à la constitution d'une culture de partage de données et la diffusion de pratiques en la matière. Le terme "données ouvertes" y a fait son apparition dès 1995 dans le secteur de la recherche environnementale, où les chercheurs ont eu besoin de partager les résultats à l'échelle mondiale<sup>7</sup>. Si les résultats de la recherche circulent depuis longtemps, la "science ouverte" se préoccupe donc aussi de celle des données sur lesquelles s'appuient ces résultats. Il s'agit d'une part de faciliter la critique des résultats publiés, mais plus encore, de fournir le matériau de futures recherches<sup>8</sup>.

<sup>4</sup> Voir :

- la circulaire du 14 février 1994 relative à la diffusion des données publiques (Journal officiel du 19 février 1994, p. 2864).

- La circulaire du 6 juillet 1994 Consultation des banques de données juridiques

- L'arrêté du 12 octobre 1994 Base de données informatisées du Journal officiel des lois/décrets (Journal officiel du 14 octobre 1994, p. 14451).

- Enfin le décret n° 94-980 du 14 novembre 1994 relatif à la délivrance de documents par le Conseil d'Etat, les cours administratives d'appel et les tribunaux administratifs (Journal officiel du 15 novembre 1994, p. 16158).]

<sup>5</sup> Décret n° 2002-1064 du 7 août 2002 relatif au service public de la diffusion du droit par l'internet

<http://www.legifrance.gouv.fr/WAspad/UnTexteDeJorf?numjo=PRMX0205836D>

<sup>6</sup> <http://www.legislation.gov.uk/>

<sup>7</sup> Notons que dans une large mesure, la critique du Groupe intergouvernemental d'experts sur l'évolution du climat (GIEC) se fonde – d'une manière certes abusive – sur la prétendue opacité des données sur lesquelles s'appuient les modèles climatologiques. Récemment deux chercheurs en appelaient à ce que l'Organisation météorologique mondiale appuie la création d'une banque communes de données climatiques

[http://www.lemonde.fr/idees/article/2009/12/21/les-donnees-climatiques-devraient-devenir-publiques-par-gil-mahe-et-alain-gioda\\_1283739\\_3232.html](http://www.lemonde.fr/idees/article/2009/12/21/les-donnees-climatiques-devraient-devenir-publiques-par-gil-mahe-et-alain-gioda_1283739_3232.html)

<sup>8</sup> Cette dynamique a donné naissance en 2006 à l'[Open Data Foundation](#).

**Dans le secteur privé**, l'échange plus ou moins ouvert de données constitue une pratique de plus en plus répandue, et se situe même au cœur du modèle économique de plusieurs géants du web.

Le pionnier en la matière est [Amazon](#) qui, dès 2005, a ouvert ses bases de données de livres via une "interface de programmation" ([API](#)<sup>9</sup>). Des professionnels du livre comme des amateurs ont pu afficher sur leur site des informations à jour sur les livres, et les ont combinées à leur propre contenu. Par ricochet, les internautes intéressés par un livre se voyaient redirigés vers Amazon pour l'acheter. Le revenu est partagé entre Amazon, et le site partenaire.

Google Maps a lui aussi proposé ses fonds de carte sous la forme d'[API](#) ouvertes : gratuites pour un site sans visée commerciale et jusqu'à un certain niveau de consultation ; payantes si le trafic est important ou pour l'accès à certaines fonctions avancées. Le cas échéant, les revenus de la publicité générée par les "Google Ads", sur les sites tiers, sont partagés entre Google et le partenaire.

A la suite de ces pionniers, la pratique s'est généralisée et fonde de nombreux modèles économiques. De nombreux sites web se construisent en assemblant des briques (données, fonds de cartes, applications...) issues d'autres sites. On désigne ces "services composites" sous le nom de "[Mashups](#)".

Il est important de noter que ce système n'est pas complètement "gagnant/gagnant" au sens où le réutilisateur est dans une dépendance réelle à l'offreur, ce dernier pouvant stopper à tout moment la mise à disposition de ses données.

A la différence du secteur public, le secteur privé est en effet entièrement libre d'ouvrir ou non l'accès à ses données. Du moins pour l'instant... La question devient en effet de plus en plus complexe lorsque les données en question proviennent, par exemple, de capteurs installés dans l'espace public, ou encore des internautes eux-mêmes : traces d'usages et de transactions, alimentation sous forme de [crowdsourcing](#) (où le site demande à ses utilisateurs de produire ou d'enrichir son contenu), publications et données personnelles sur les réseaux sociaux. *"Statistiques de fréquentations des sites web, composition du graphe social, recommandations dans son réseau social, achats en ligne et composition du panier, déplacements dans la ville identifiés par le mobile, consommation électrique identifiée par les compteurs intelligents... chacun d'entre nous génère au quotidien une multitude de traces numériques qui sont autant de données susceptibles d'être réexploitées par des tiers, et de circuler [...]. Les sites web ont-ils le droit de commercialiser ces données générées par leurs utilisateurs ?"* questionne la chercheuse Valérie Peugeot<sup>10</sup>. Les réponses à cette question seront sans doute plurielles...

## Vers le "web des données"

L'évolution des technologies de l'information et de la communication, particulièrement des formats et des capacités d'échange de contenus, va de pair avec le mouvement d'ouverture et de réutilisation des données. Elle le permet autant qu'elle le découpe. Le web devient l'espace de jeu et de partage par excellence.

Le web connectait au départ des documents, des "pages". Désormais, il interconnecte des applications, des services et des informations de base, les données. Un document avait son adresse (l'"URL") et ses descripteurs (un titre, un auteur, une date de création, des mots-clés), il en va désormais de même des données. Grâce au web, un site web pourra par exemple agréger en un seul endroit une multitude d'informations sur une ville : des cartes, la localisation de services ou de lieux publics, de l'information trafic, la météo, des statistiques économiques ou sociales, etc.

Pour qualifier l'importance qu'est en train de prendre la donnée dans le web d'aujourd'hui et de demain, Tim O'Reilly et John Batelle, les inventeurs du concept de [web2.0](#), ont proposé en 2009 le terme de [web2](#) ("web au carré") : par là ils pointaient l'idée que le web, en raison de la quantité de données - publiques, privées, de bien commun, statistiques, géographiques, personnelles... et des

<sup>9</sup> Voir lexique. Amazon ne livre donc pas l'ensemble de sa base de données : elle permet à des programmes informatiques de l'interroger en ligne, sous la forme de "requêtes" auxquelles la base de données répond avec des fiches (ouvrages et autres produits).

<sup>10</sup> in Peugeot 2010 p.7

possibilités de [mashup](#), ne se développait plus de manière arithmétique mais exponentielle. C'est une puissance supplémentaire du web et d'une capacité nouvelle pour l'homme de combiner des informations.

### "RAW DATA NOW" LES DONNEES BRUTES MAINTENANT !

[Intervention de Tim Berners-Lee lors de la conférence TED](#) (2009)

TALKS

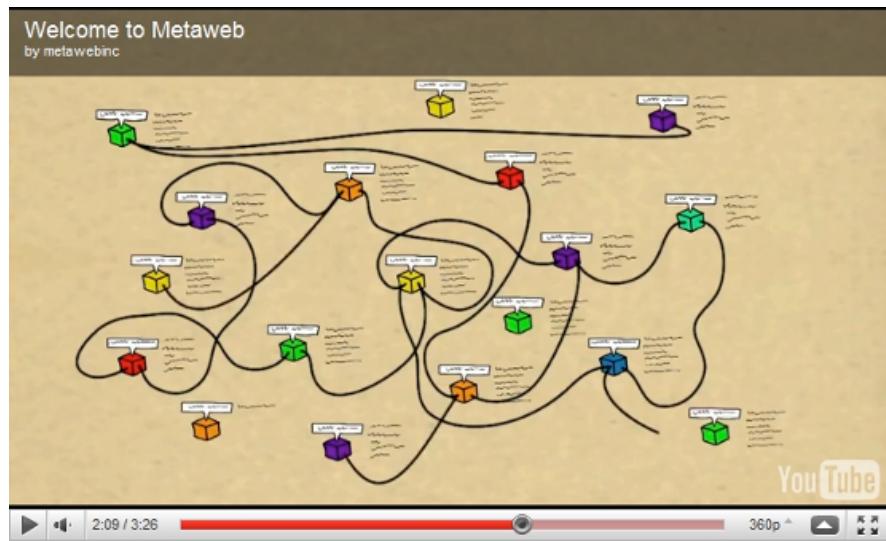
### Tim Berners-Lee on the next Web

TED2009, Filmed Feb 2009; Posted Mar 2009



Lors de sa présentation à la conférence TED en mars 2009, Tim Berners-Lee, l'inventeur du World Wide Web, en appelle au partage et à l'ouverture des données brutes. "Grâce au "www", on a pu mettre en ligne sur le web des milliers de documents, explique-t-il. Maintenant il faut y mettre des données". Quelles différences y a-t-il entre documents et données ? Les documents sont à lire, à analyser. Les données, si elles sont dans un format [interopérable](#), sont directement utilisables, combinables... C'est le principe du "[LinkedData](#)" ("données liées") : les données du web de demain seront des objets relationnels.

La profusion de données et d'informations sur le web, et leur interrelation requièrent de penser de nouveaux formats de structuration, de recherche, de lien entre celles-ci. La description de ces relations devient le moteur des applications et services de demain. C'est là le domaine de recherche du web sémantique. Il ne s'agit plus de créer des liens entre des mots isolés, mais entre des entités sémantiques, comme l'explique cette vidéo proposée par Metaweb, un éditeur de bases de données ouvertes (Freebase).



# 1. OUVRIR ET UTILISER DES DONNEES PUBLIQUES TERRITORIALES

## 1.1. UNE DONNÉE PUBLIQUE, C'EST QUOI AU JUSTE ?

Comme toute organisation aujourd'hui, un acteur public utilise l'informatique pour préparer toutes ses décisions, produire tous ses services, évaluer toutes ses actions. Et l'informatique traite des données : par exemple, des fonds de cartes et des informations attachées aux cartes, des statistiques, des descriptions de services et de lieux publics, des mesures, des études et rapports, des barèmes, des textes réglementaires, des informations temps réel, et de bien d'autres choses. Ces informations sont la matière première de l'action publique. Mais elles pourraient aussi bénéficier à d'autres acteurs publics, ainsi qu'à des entreprises innovantes, des associations, des chercheurs, des citoyens. C'est l'idée qui sous-tend la directive européenne sur la "réutilisation des données publiques" (2003), transcrite en droit français en 2005 sous la forme d'une révision de la loi du 17 juillet 1978 sur l'"amélioration des relations entre l'administration et le public" : les "données publiques", financées par l'impôt, doivent pouvoir être réutilisées par d'autres acteurs, au service de la qualité des services sur le territoire, de la croissance économique, de la connaissance et du débat démocratique.

### La définition juridique

En France, la loi considère la notion d'information publique, qui inclut les données, comme une information *produite ou reçue* "dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission" [article 1 de la loi de 1978].

C'est donc la finalité de l'action qui détermine le caractère public d'une information : les données produites par un opérateur privé dans le cadre d'une mission de service public, sont des données publiques. En revanche, les données qui participent à une mission de service public industriel et commercial ne sont pas *a priori* "publiques".

La loi française établit donc une obligation générale d'ouverture des données publiques, tout en prévoyant quelques exceptions :

- les données relevant de la vie privée,
- les données relevant de la sécurité nationale,
- les informations présentes dans des documents sur lesquels des tiers détiennent des droits de propriété intellectuelle.

Toutes les autres données publiques sont réutilisables de droit. La loi française est même allé plus loin que la directive européenne en créant un droit opposable à la réutilisation de ces données. Une administration ne peut s'opposer à la demande d'un acteur privé, quel qu'il soit, de voir ses données réutilisées. En cas de refus ou de non réponse, tout acteur privé, individu ou organisation, peut saisir la [CADA](#) dont le rôle est de "veiller à la transparence de l'action administrative et se prononcer sur les modalités de réutilisation des informations publiques" et qui constitue une véritable voie de recours précontentieuse<sup>11</sup>.

### Le cas des données nominatives

Les données nominatives peuvent être des données publiques lorsqu'elles ne relèvent pas de la vie privée : par exemple les données qui concernent les personnes publiques dans l'exercice de leur fonction ; ou encore les données concernant des personnes qui n'ont pas de responsabilité publique, lorsque ces personnes interviennent dans le cadre d'une activité publique. En revanche, tout en étant collectées dans le cadre d'une mission publique, les données du recensement, par exemple, ne sont pas publiques.

<sup>11</sup> <http://www.cada.fr/fr/presentation/center2.htm>

Le droit distingue également des cas où une donnée nominative peut avoir une publicité limitée : par exemple l'annonce d'un mariage est publié en mairie mais ne peut faire l'objet d'une publication sur l'internet ; au titre du droit à l'oubli, il n'est pas permis de publier indéfiniment sur l'internet des jugements, impliquant donc des personnes, alors que cette information est publiée localement.

La publication de données nominatives doit donc faire l'objet d'une analyse juridique détaillée pour s'assurer qu'elle entre bien dans le cadre de la loi.

### **L'anonymisation des données**

Il peut être intéressant de rendre publiques des données nominatives en les anonymisant. Le secteur de la santé, par exemple, fait un usage abondant de données anonymisées, dont certaines sont publiques (progression de la grippe, etc.). L'anonymisation est cependant un travail très délicat. Certains chercheurs avancent que l'anonymisation absolue est impossible. Là encore, il existe en France une expertise sur ce sujet, qu'il convient de consulter<sup>12</sup>.

---

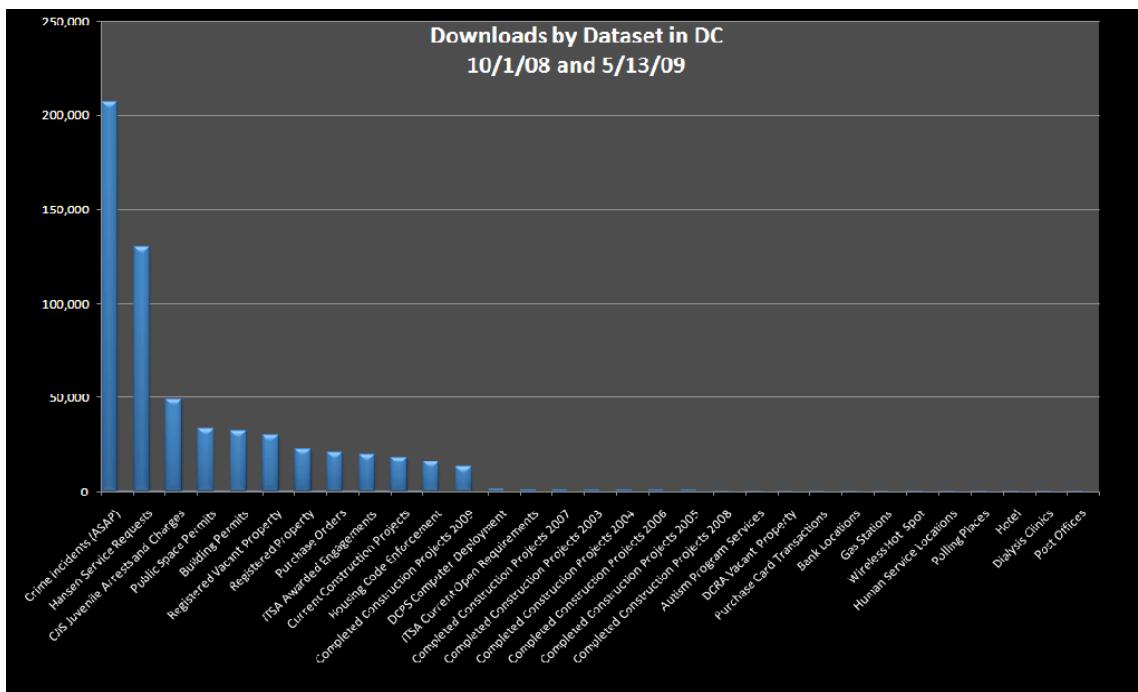
<sup>12</sup> Livre blanc "Anonymisation des Données à caractère personnel", écrit par Bruno Rasle:  
<http://www.cortina.fr/livres-blancs.php>

## **1.2. QU'ATTEND-ON DE L'OUVERTURE ET LA REUTILISATION DES DONNEES PUBLIQUES ?**

Les motivations qui prévalent à l'ouverture comme à la réutilisation des données publiques peuvent être très diverses, selon le point de vue du protagoniste sous lequel on se place : producteurs, réutilisateurs. Dans les deux cas, les attentes de "création de valeurs" sont très fortes que celles-ci aillent dans le sens d'une plus grande innovation sociale, d'un nouveau dynamisme économique, d'une plus grande efficacité des procédures administratives, d'une amélioration des services au public, ou encore d'un nouveau contrat démocratique. En tout cas, la valeur attendue ne serait se restreindre à la seule valeur économique.

Cela étant, la mise à disposition des données ouvre par définition un espace d'incertitude : l'administration qui "libère" ses données ne peut pas tout savoir, ni tout contrôler, de ce qui en sera fait par les "réutilisateurs". Le mouvement étant par ailleurs relativement récent, les acteurs susceptibles de tirer parti des données publiques ne se sont pas encore tous manifestés.

Ainsi, lors d'une première analyse de la réutilisation des données du portail américain [data.gov](http://data.gov), on a pu être surpris de constater qu'en 2008 et 2009, l'exploitation des données criminelles a concentré à elles seules 80% des réutilisations<sup>13</sup>.



D'où la nécessité d'analyser et de comprendre au mieux, au regard des initiatives existantes, les objectifs attendus et les impacts réels. Cette partie s'attache à présenter les principaux objectifs avancés par les parties prenantes, et de les mettre en regard des résultats déjà observables. Malgré le manque de recul dans le temps, il est indispensable de partager largement les initiatives et de se livrer au fur et à mesure à leur évaluation.

### **1.2.1. Efficacité administrative et innovation publique**

Améliorer les outils de travail internes et moderniser les processus bureaucratiques sont des objectifs largement avancés par les acteurs publics : gagner en efficacité, réduire les redondances, améliorer la gestion des connaissances, mieux évaluer l'action publique, réduire les coûts...

<sup>13</sup> Hubert Guillaud "Mesurer les effets de la libéralisation des données"

<http://www.internetactu.net/2009/09/16/critiquesdu-web%C2%B2-24-les-effets-de-la-liberation-des-donnees/>

Les services des administrations produisent - seules ou en se faisant aider par des tiers - des informations destinées à éclairer les propositions d'actions publique et les décisions des élus (statistiques, diagnostics, données prospectives, rapports d'évaluation, etc.). Ces informations restent la plupart du temps au sein des services, au mieux sont-elles stockées sur un serveur partagé - le fameux "intranet" -, et référencées par des mots clés. La circulation de ces informations n'est pas chose aisée d'un service à l'autre, et encore moins d'une administration à une autre, le plus souvent par méconnaissance et manque de visibilité.

#### ► **Décloisonner l'action publique :**

Les divers ministères ou directions d'une collectivité travaillent bien souvent sur des objets communs, mais elles ont malgré tout tendance à les traiter chacune de son côté, à reproduire chacune les mêmes bases de données (statistiques, géographiques, etc.). Si chaque direction faisait savoir aux autres de quelles données elle dispose, différentes possibilités s'ouvriraient :

- Limiter les redondances : une seule base de données pour une même fonction, alimentée et mise à jour par tous ceux qui y ont recours ;
- Croiser, comparer, combiner des informations : les informations des directions en charge, par exemple, de la voirie, des transports, de l'environnement, ont tout intérêt à être croisées. D'autres services peuvent également en avoir besoin, qu'il s'agisse des "bureaux des temps" ou des agences de développement économiques chargées d'étudier des programmes d'implantation d'entreprises. Ces croisements se font déjà aujourd'hui, mais de manière *ad hoc* et complexe. Ils pourraient devenir normaux et quotidiens.
- Mieux évaluer l'action publique : la plupart des décisions publiques, notamment locales, ont des conséquences dans un grand nombre de domaines. Leur efficacité ne peut s'évaluer dans le cadre du seul département qui en a la charge.

#### ► **Rendre l'action publique plus lisible**

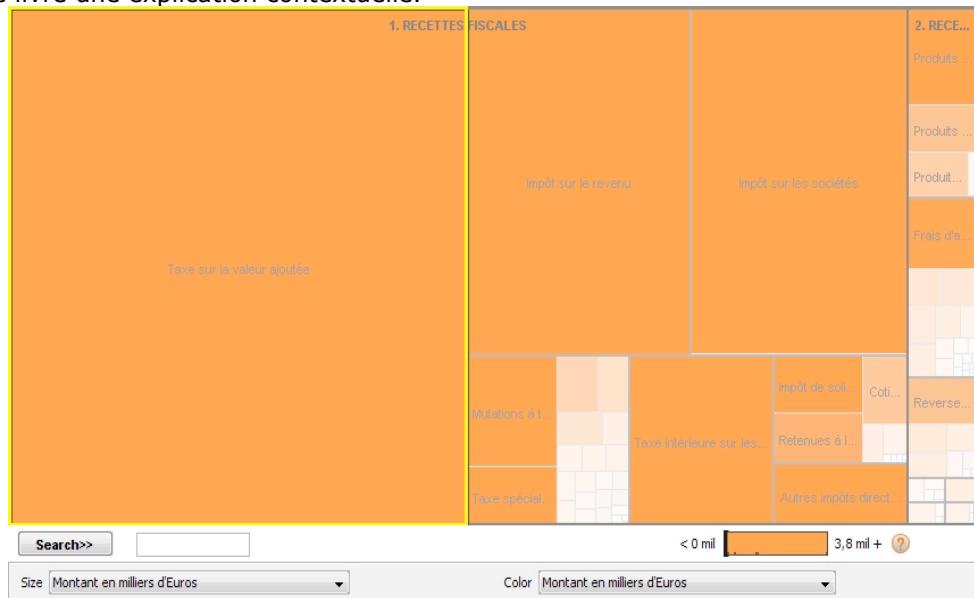
La mise à disposition des données publiques est une manière – certes indirecte – de rendre l'action publique plus intelligible, dans ses processus, ses choix et ses conséquences. Il n'est pas facile de la comprendre lorsque celle-ci s'exprime dans des longs documents ou dans un langage administratif certes nécessaire à la précision juridique, mais peu parlant pour les citoyens.

Ainsi, de très nombreux travaux portent aujourd'hui sur la "visualisation de données". Il s'agit de traiter un vaste ensemble d'informations, parfois très hétérogènes, pour le représenter de manière visuelle et le rendre plus compréhensible, voir pour détecter des corrélations jusqu'alors invisibles.

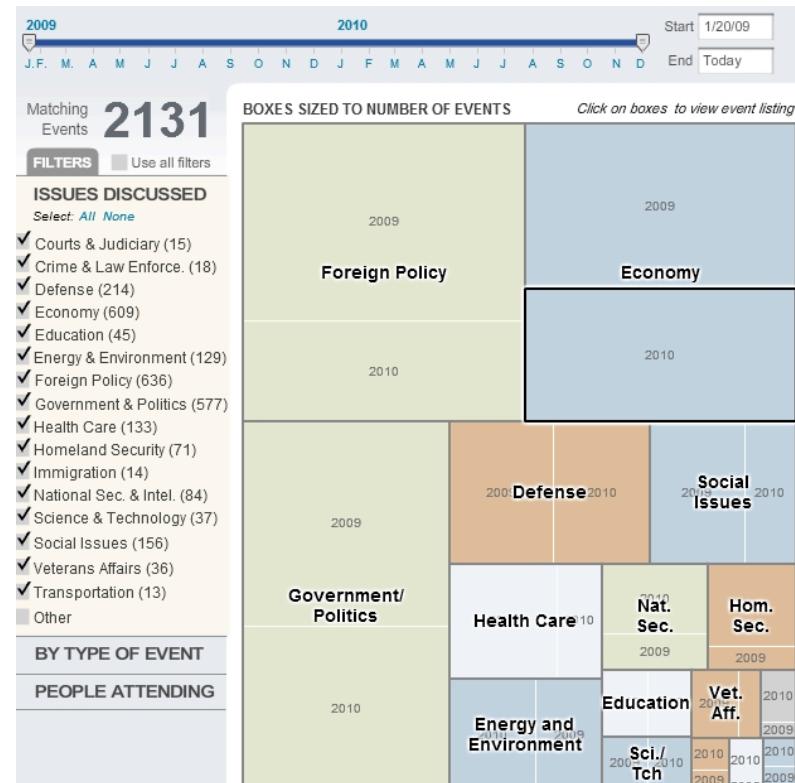
Le [Treemapping](#) fait partie de ces nouvelles formes de visualisation, mettant en exergue les structures hiérarchiques des informations contenues.

## QUELQUES EXEMPLES DE TREEMAPPING

Les chercheurs d'IBM auteurs d'un ensemble d'outils de visualisation censés être accessibles à tous, ManyEyes, ont "cartographié" les dépenses du budget français<sup>14</sup>. Un clic sur chaque grand ensemble livre une explication contextuelle.



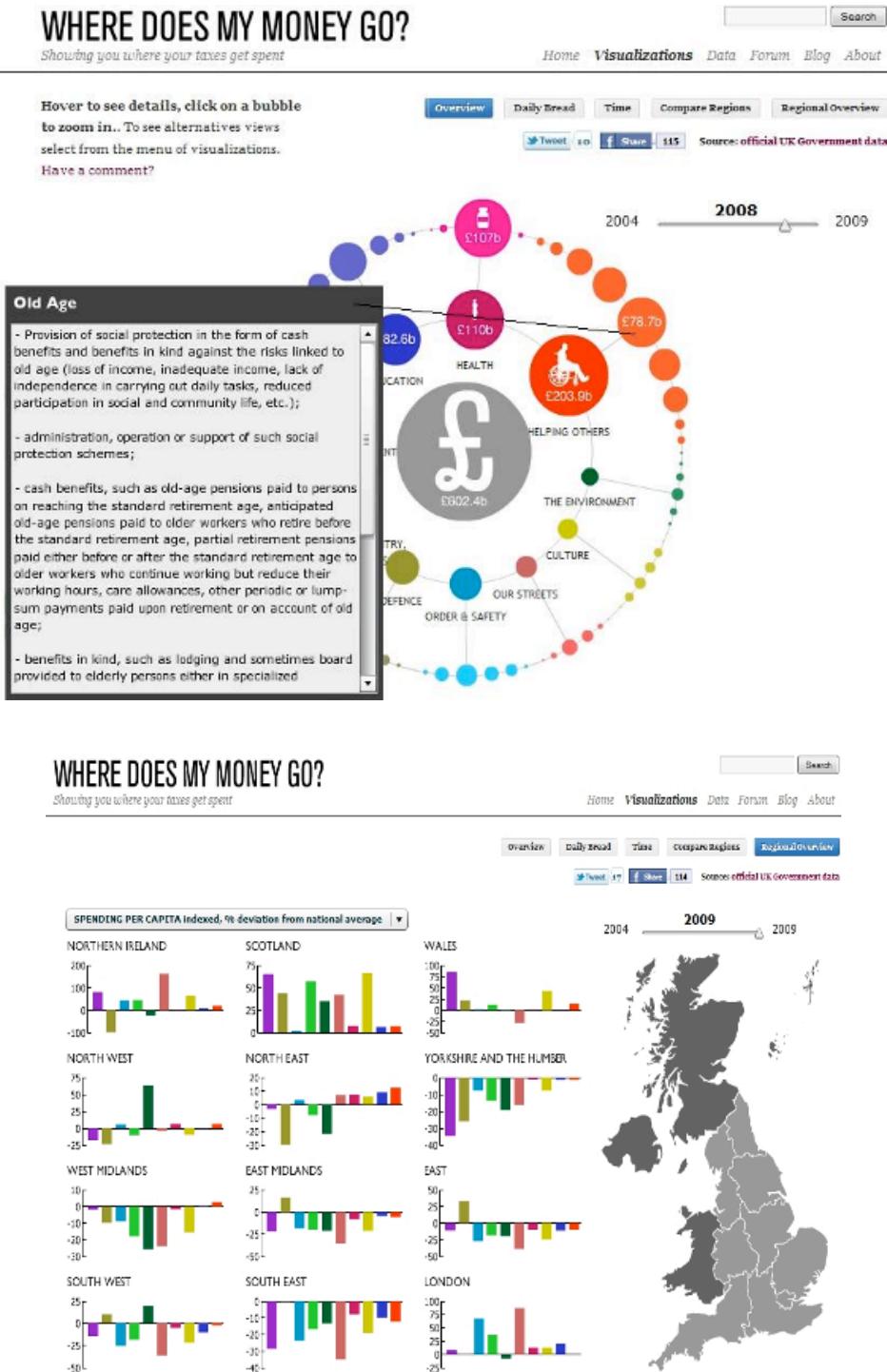
Le Washington Post<sup>15</sup> s'est également servi du Treemapping pour donner à voir différemment l'agenda du Président Obama. En faisant ressortir visuellement les thèmes abordés lors des différentes réunions, on donne à lire les sujets qui marquent l'intérêt du Président, ou qui contraignent l'actualité politique.



<sup>14</sup> <http://www-958.ibm.com/software/data/cognos/maneyes/visualizations/recettes-de-letat-francais-2007-pr>

<sup>15</sup> [http://projects.washingtonpost.com/potus-tracker/?wpisrc=nl\\_fed](http://projects.washingtonpost.com/potus-tracker/?wpisrc=nl_fed)

En Grande-Bretagne, le site "[Where does my money go?](#)" rend visible d'une autre manière la répartition du budget national, ainsi que des différentes régions, avec des comparatifs possibles.



### ► Inviter au partage et à la collaboration

En ouvrant ses données, l'administration invite de fait d'autres acteurs, ainsi que les citoyens, à contribuer à leur tour à enrichir une sorte de "domaine public" des informations relatives à un territoire ou un domaine.

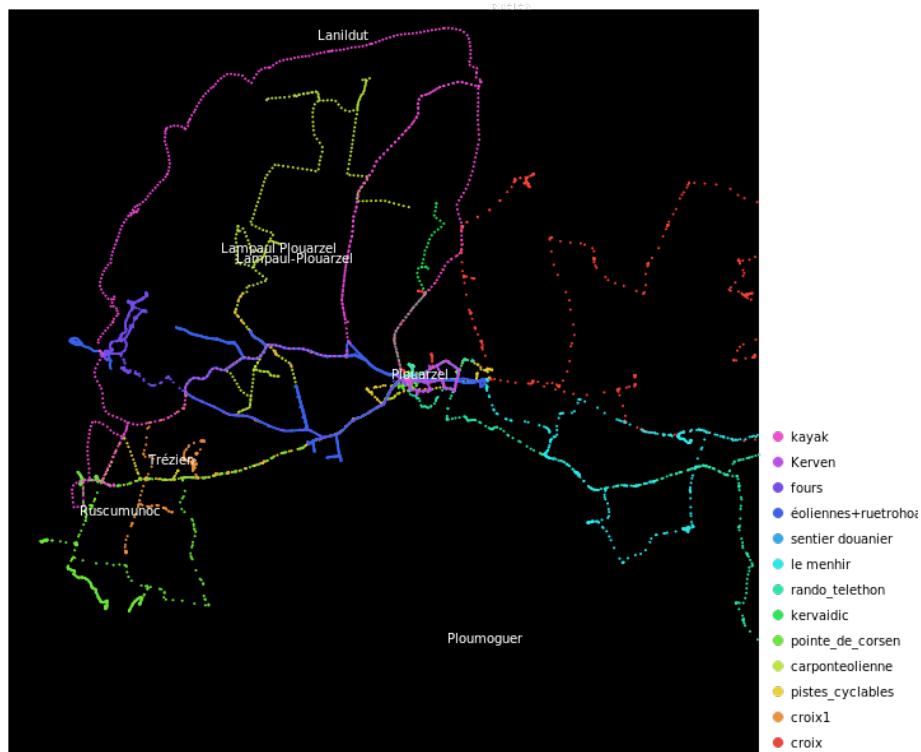
Ces acteurs peuvent d'abord être des entreprises, à commencer bien sûr par celles qui ont la charge de services indispensables sur un territoire. On sait, par exemple, combien la réticence de tous les acteurs de la mobilité à partager leurs informations rend difficile, presque partout en

France, la mise en place de services d'information multimodale sur les transports, qui seraient pourtant si importants pour la vie quotidienne des habitants (et pour l'environnement). Pensons également aux trésors d'information que détiennent les réseaux d'énergie et d'eau, les entreprises chargées du traitement des déchets, les opérateurs de télécommunication, la grande distribution, les banques, La Poste... Plusieurs de ces entreprises examinent aujourd'hui quelles informations elles pourraient partager sans dévoiler, ni la vie privée de leurs clients, ni leurs secrets stratégiques. Si elles le font, c'est bien parce que les acteurs publics ont pris le devant.

Les citoyens peuvent également contribuer à enrichir les données "publiques". A Brest, Rennes, dans la petite commune de Plouarzel (Finistère), mais aussi en Haïti juste après le tremblement de terre, la population (et dans le cas d'Haïti, les internautes du monde entier) a été appelée à compléter, corriger, enrichir les cartes du territoire. Les données disponibles, généralement issues du Système d'information géographique (SIG), étaient mises à disposition du site de cartographie collaborative OpenStreetMap et de sa communauté. Et celle-ci l'a rendu au centuple à la collectivité. Certains ont précisé le tracé des rues ou localisé les points d'intérêt. D'autres ont cartographié l'accessibilité du territoire aux handicapés. D'autres ont tracé les itinéraires cyclables, etc.

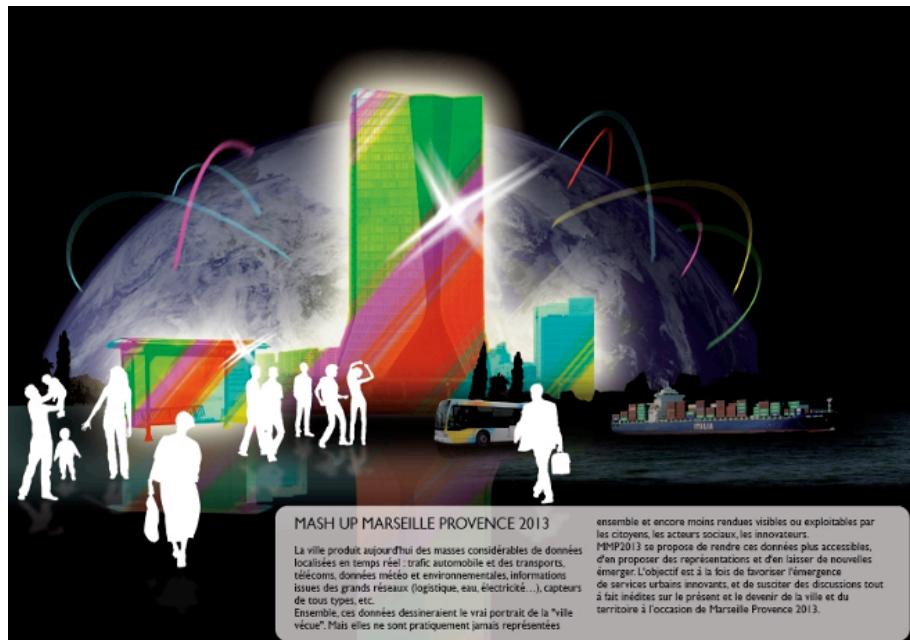
### PLOUARZEL ET LE CROWDSOURCING GRANDEUR NATURE

Confrontée au problème de mise à jour et d'édition de son plan communal, et ne disposant pas des moyens nécessaires pour faire fonctionner un service technique, cette commune de 3150 habitants a fait le choix de mettre les cartes qu'elle avait à disposition sur OpenStreetMap, et de proposer à la population de les enrichir. Au cours d'une "cartopartie", coordonnée par le centre multimédia de Plouarzel, une quarantaine d'habitants volontaires, équipés de GPS par la mairie, ont sillonné le territoire "à pieds, à cheval, à vélo, à moto, en voiture... et même en kayak<sup>16</sup> pour en répertorier les sites les plus remarquables et les inscrire sur les cartes d'OpenStreetMap. Ainsi enrichies, celles-ci sont devenues aujourd'hui l'outil de référence de l'Office municipal de Tourisme. Cette démarche participative, très innovante, a attiré sur la commune une notoriété nouvelle, sans compter le sentiment que les habitants peuvent ressentir d'avoir contribué au bien public.



<sup>16</sup> <http://wiki.openstreetmap.org/wiki/Plouarzel>

## MASHUP MARSEILLE PROVENCE 2013



Marseille et les villes environnantes seront en 2013 la "capitale culturelle de l'Europe". Et si ce grand projet dont le territoire espère beaucoup était produit en commun avec les habitants, les associations et les entreprises qui y vivent ? C'est l'idée qui sous-tend le projet "MMP2013". L'idée consiste à verser un vaste ensemble de données urbaines, publiques, privées et citoyennes, dans un même espace à partir desquelles elles pourraient être réutilisées de diverses manières. Et pour en faciliter l'accès, le projet prévoit de partager également des outils d'analyse et de visualisation de ces informations.



17

<sup>17</sup> Illustration Anaïs Triolaire et Julien Defais - atelier Marseille Provence 2013

## 1.2.2. Transparence de l'action publique et vie démocratique

### ► Dédramatiser la "transparence"

A l'évidence, l'ouverture des données publiques est un outil de transparence. Elle permet à des tiers, quelles que soient leurs intentions, d'examiner certains pans de l'action publique et s'ils le jugent pertinent, d'en dénoncer les dysfonctionnements.

La transparence a aujourd'hui mauvaise presse. Elle s'assimile volontiers au soupçon systématique, à la dénonciation, ou aux dérives numériques dont l'affaire des télegrammes diplomatiques américains révélés par WikiLeaks est le plus récent symbole.

Les pays nordiques et anglo-saxons en font pourtant une pierre angulaire de leur vie démocratique. En Suède, une loi datant de 1776 autorise tout citoyen à demander auprès des autorités concernées en moins de vingt quatre heures, d'avoir accès aux notes de frais d'un Ministre<sup>18</sup>. En Angleterre, la culture de la transparence est aussi très forte. En novembre 2010, le gouvernement de David Cameron annonçait ainsi que toutes les dépenses administratives dépassant 25 000 £ (environ 29 500 €) seront publiées en ligne, ainsi que l'intégralité des contrats correspondants, soit presque 200 000 lignes articles de dépenses. Bien sûr l'objectif n'est pas seulement de vivifier le débat démocratique, il est également financier : faire pression sur les tarifs de certaines prestations en faisant jouer la concurrence<sup>19</sup>, et ainsi mieux employer l'argent public ("modern deregulation"). Soit une plus grande transparence au service d'une plus grande efficacité de l'action publique.

Mais la transparence est également une règle démocratique partout ailleurs, notamment en France. Elle est déjà au cœur de la loi de 1978, qui instaure la "liberté d'accès aux documents administratifs", et c'est cette loi dont la révision en 2005 encadre désormais aussi la mise à disposition des données publiques à des fins de réutilisation.

C'est au nom de cette transparence que l'association [Regards Citoyens](#) a créé le site <http://www.nosdeputes.fr/>, qui extrait de plusieurs sources (en général de pages web que des logiciels "lisent" pour en extraire les données présentes) de nombreuses informations sur chaque député : les commissions auxquelles il appartient, ses votes, ses interventions, sa présence ou son absence en séance, etc. Voici comment le site se présente aux internautes : *"En synthétisant les différentes activités législatives et de contrôle du gouvernement des élus de la nation, ce site essaie de donner aux citoyens de nouveaux outils pour comprendre et analyser le travail de leurs représentants. Conçu comme une plateforme de médiation entre citoyens et députés, le site propose à chacun de participer et de s'exprimer sur les débats parlementaires. Au travers de leurs commentaires, les utilisateurs sont invités à créer le débat en partageant leur expertise lorsque cela leur semble utile. Peut-être pourront-ils ainsi nourrir le travail de leurs élus ?"* Certes, le site permet de voir quels députés détiennent les records d'absence en séance, mais il rend aussi visible l'activité de la masse des députés qui travaillent, et contribue probablement plus à revaloriser le travail politique qu'à l'inverse.

Nosdeputes.fr a été créé par une association parce que le Parlement ne proposait pas l'accès à ces données, pourtant éminemment "publiques" par vocation. C'est dire l'importance, pour les acteurs publics, de prendre les devants : s'ils prennent l'initiative de rendre visible les informations qui permettent le contrôle démocratique, ils rendront le soupçon sans objet et permettront que s'engage un dialogue constructif sur le pourquoi de tel choix, de telle dépense. Dans le cas contraire, ils s'exposent à ce que les données soient malgré tout divulguées, souvent par des groupes qui poursuivent un intérêt particulier, éventuellement entachées d'erreurs, et dans un contexte où ils auront à se défendre non seulement sur le fond, mais sur le pourquoi de leur opacité.

Plus la transparence est choisie, plus ses nécessaires limites – comme on le voit s'agissant des télegrammes diplomatiques, dont même WikiLeaks a expurgé ceux qui faisaient courir des risques aux personnes – peuvent faire l'objet d'une discussion dépassionnée.

<sup>18</sup> <http://www.sweden.gov.se/sb/d/2184/a/15521> A noter que, en cas d'ouverture d'une instruction judiciaire, ou de poursuites quelconques, la loi protège les fonctionnaires suédois qui délivrent les documents.

<sup>19</sup> <http://www.guardian.co.uk/politics/2010/nov/19/government-data-accounts-private-firms>

## ► La participation sur des bases communes

On connaît l'importance, et les limites, de la "démocratie participative". Trop souvent, les débats citoyens qu'organisent les collectivités rassemblent des audiences clairsemées, composées avant tout de "citoyens professionnels", d'habitues du débat public qui ont certes acquis une réelle expertise, mais qui, du coup, ne représentent pas plus (et parfois moins) que les élus les "citoyens". Les débats publics en ligne sont moins aisément monopolisés par ces acteurs par ailleurs nécessaires, mais il est également difficile de savoir qui y intervient, et le profil socio-culturel de ceux qui y participe est souvent très typé.

L'ouverture des données publiques ne résout pas magiquement ces problèmes, mais elle ouvre une possibilité nouvelle, et déterminante : celle de partager, entre toutes les parties du débat, une base d'informations, de représentations et de simulation commune.

Supposons que les acteurs publics, les citoyens, les entreprises concernées, les associations, travaillent tous ensemble sur les mêmes informations ; qu'ils partagent des outils d'analyse, voire de simulation (modèles économiques, cartographie pour des projets d'urbanisme, etc.) ; que chacun puisse proposer ses analyses et permettre aux autres de les discuter en les comprenant... ne peut-on imaginer que la "démocratie participative" y trouverait une seconde jeunesse ?

### **1.2.3. De nouveaux services, des services améliorés pour la vie quotidienne**

#### ► **Innovation sociale et innovation économique**

Face à la l'urgence et la complexité des enjeux territoriaux, aux contraintes financières et à l'émergence de nouveaux défis (la ville durable, le vieillissement, le lien social...), à la diversité des besoins, les acteurs publics ne peuvent plus être seuls à délivrer (ou déléguer, ce qui revient presque au même) les services essentiels à leurs habitants. Il faut décloisonner, faire venir et associer en permanence des idées et des énergies neuves. Une première réponse a été cherchée dans les "partenariats public-privé", mais ceux-ci concernent le plus souvent de très grandes entreprises, chargées de produire un service public précisément défini et, le plus souvent, choisies sur des critères avant tout financiers.

En ouvrant les données utiles à la production de services sur son territoire, la collectivité publique invite d'autres acteurs à intervenir : des associations locales qui connaissent bien les besoins d'une population, des médias régionaux, des jeunes entreprises inventives, des citoyens malins...

#### **LES DONNEES DE TRANSPORT DE RENNES METROPOLE EN ACCES LIBRE**

A Rennes, l'ouverture des données de transport, et particulièrement celles liées à l'utilisation de l'infrastructure VeloStar, a suscité une floraison d'applications et services nouveaux : des alertes trafic en temps réel, un état d'utilisation des vélos dans les différentes stations, un état d'occupation des parcs relais, une géolocalisation des points de vente, un état des fonctionnements des ascenseurs et des escalators, les horaires des bus et des métros...

The screenshot shows the homepage of keolis-rennes.com. At the top, there's a blue header bar with the text 'keolis-rennes.com' and 'DATA BETA'. Below the header, a sub-header reads 'LES DONNÉES TRANSPORT DE RENNES MÉTROPOLE EN ACCÈS LIBRE'. A navigation menu at the bottom of the header includes 'Accueil', 'Le projet', 'FAQ', 'Forum', 'Les données' (which is highlighted in blue), and 'Contact'. The main content area has a heading 'Bienvenue sur l'entrepot des données transport de Rennes Métropole !'. It discusses the opening of STAR network data and the Le vélo STAR service, noting it's a first in France. It aims to facilitate development of applications and services. It also mentions a 'concours' (competition) for developing new services. The page features icons for developers and users, and a sidebar with a 'CONCOURS' banner.

Seuls, l'opérateur privé ou l'acteur public n'auraient peut-être pas eu ces idées-là, ou n'auraient pas eu les ressources internes pour les développer. Ces services ont incontestablement contribué au succès de VeloStar et à son adoption rapide par la population rennaise.

Ouvrir des données – par exemple les données des mobilités-, c'est permettre à d'autres acteurs de développer de nouveaux services, ou de faire évoluer ceux qui existent : quand un jeune programmeur crée autour des vélos en libre-service une application mobile qui permet de s'informer de leur disponibilité à une station donnée, il contribue si nettement au succès de l'opération... que l'opérateur des vélos a tout fait pour s'arroger *a posteriori* l'exclusivité de telles applications, oubliant que sans innovateurs extérieurs, il n'aurait peut-être jamais pensé à proposer une telle possibilité. Dans d'autres villes, on voit que c'est par des acteurs extérieurs que

des esquisse des services d'information multimodale sur les transports émergent. Ailleurs, des entrepreneurs associent covoiturage et transports publics. Comment un seul acteur en charge des transports, si intelligent et bien financé soit-il, pourrait-il faire aussi bien et aussi vite ?

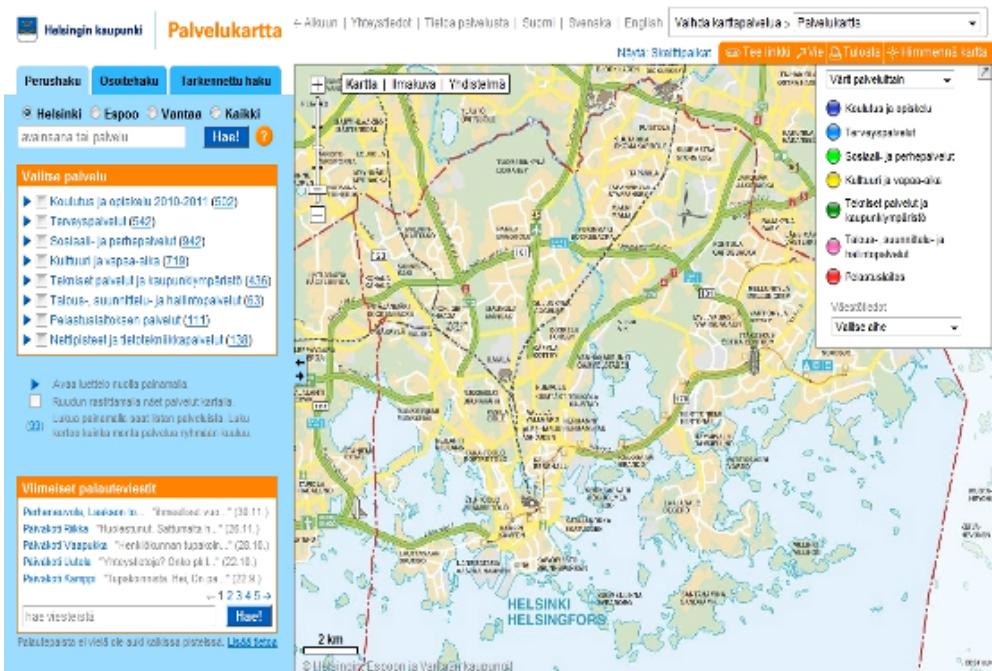
Innovation sociale, innovation publique et innovation économique se rejoignent. Les données forment la matière première des trois. Quelles que soient leurs motivations, des acteurs se saisiront des données mises à disposition pour proposer des services. Certains innoveront en matière technique, d'autres par la pertinence de leur réponse à certains besoins, d'autres par leur proximité à une population donnée – et certains échoueront aussi. Fondés sur des informations communes, ces services pourront plus aisément collaborer, entre eux comme avec les services que la collectivité continuera de fournir, pour des raisons évidentes de continuité et d'égalité.

En définitive, sans presque rien débourser, les acteurs publics peuvent améliorer la vie quotidienne de leurs administrés, en se contentant de libérer les énergies latentes et de les laisser s'exprimer. C'est, pour l'acteur public, un moyen potentiel de dynamisation et d'élargissement de la fourniture de services aux citoyens. C'est aussi un moyen d'inviter la population à s'engager, à entreprendre ou au moins, à prendre en mains certains de ses problèmes quotidiens.

### APPS FOR DEMOCRACY

Le [Forum Virium](#), "laboratoire vivant" d'innovation qui réunit entreprises et acteurs publics d'Helsinki (Finlande), organise chaque année un concours d'applications fondées sur les données publiques. S'il a repris le nom du concours équivalent lancé à Washington D.C. : "Apps for Democracy" (des applications pour la démocratie), les applications qu'il suscite se situent beaucoup plus du côté de la vie quotidienne que de la discussion démocratique. Le lauréat de 2010 est ainsi un outil mobile qui permet de scanner le code barre d'un livre pour savoir s'il est disponible dans une des bibliothèques municipales.

L'avantage de ces formes de stimulation est aussi d'obtenir des applications peu chères et souvent originales. Le site [Palvelukartta](#), qui dresse la carte des services publics de la ville d'Helsinki, est ainsi devenu le meilleur annuaire de la ville. Tout en permettant aux différents acteurs publics de la ville de mieux communiquer entre elles en disposant enfin d'un annuaire fiable, il permet également de voir les carences des services publics sur certains quartiers. Au sein des services publics, ces formes d'animation jouent aussi un rôle d'émulation : L'association a contribué à libérer plus de 200 bases de données de la ville, parmi lesquels une centaine de systèmes d'informations géographiques différents<sup>20</sup>...



<sup>20</sup> <http://www.internetactu.net/2010/07/13/donnees-publiques-ouvertes-comment-faire/>

## ► Un effet sur l'attractivité des territoires

La mise à disposition d'un plus grand nombre de données économiques, touristiques, ou liées à la vie quotidienne, peut renforcer l'attractivité d'un territoire.

Les acteurs du tourisme peuvent proposer aux publics qu'ils connaissent bien des services plus complets, des informations plus riches – et peuvent contribuer à en traduire une plus grande partie, et dans d'autres langues, que ce que des organismes publics aux moyens limités, proposent.

Des entreprises cherchant à s'implanter, des familles cherchant à se loger, peuvent disposer d'une connaissance plus fine du tissu économique, d'une meilleure vision des services accessibles ou de la vitalité de la vie locale.

L'ouverture en elle-même, et son effet mobilisateur, sont déjà des moyens d'animer une dynamique positive rejaillissant sur l'image de la collectivité, de son territoire. Les démarches ludiques - concours, etc. - de stimulation de la réutilisation des données, ou d'enrichissement participatif de celles-ci constituent à elles seules de nouveaux modes de relations aux publics locaux comme extérieurs au territoire.

## ► Des bénéfices nets, des défis complexes

L'idée que des acteurs entièrement indépendants de l'administration, sans autre lien avec elle que la signature (éventuellement électronique) d'une licence de réutilisation de données, contribuent à produire et délivrer des services au public, est inhabituelle. Séduisante, voire puissante, elle pose néanmoins des questions qu'il ne faudra pas négliger à terme : quelle responsabilité l'acteur public partage-t-il avec les opérateurs de ces nouveaux services ? Quelles nouvelles exigences cela fait-il peser, de fait, sur la qualité ou la fraîcheur de ces données ? Que faudra-t-il prévoir dans les licences en matière de transformation des données, de délais de redistribution (pour les données temps réel), d'indication de la source publique, de responsabilité ?... Et enfin, l'émergence de ces services est-elle de nature à concurrencer le cœur des missions publiques – et dans ce cas, quelles en sont les conséquences ?

Cependant, ces questions ne peuvent pas s'appréhender de manière théorique. Les réponses n'émergeront que de manière progressive, à l'expérience.

### 1.2.4. Valoriser le patrimoine immatériel local

## ► Mettre en valeur les données culturelles, patrimoniales

Comme beaucoup d'organisations, les collectivités publiques exploitent souvent assez mal les trésors d'informations qu'elles détiennent – faute de moyens, d'intérêt, de demande, de compétences disponibles... C'est le cas, bien souvent, des "données culturelles". Certaines collectivités territoriales détiennent des fonds documentaires, patrimoniaux, considérables (documents iconographiques, œuvres numérisées, affiches, documents audiovisuels, sonores, topographiques, etc.). Ces fonds sont le plus souvent disponibles dans des bibliothèques ou des archives, parfois exposés, mais en général méconnus et sous-utilisés.

Faire le choix de les rendre accessibles en ligne, et réutilisables, selon un cadre de réutilisation donné, peut considérablement augmenter la visibilité de ces ressources et, par ricochet, contribuer à la notoriété du territoire, voire à faire émerger de nouvelles activités culturelles. Des généalogistes aux historiens, des experts aux enseignants, des amateurs de vieilles photos aux communautés d'"anciens" désireux de transmettre leur histoire aux plus jeunes<sup>21</sup>, les raisons de mieux exploiter ces fonds ne manquent pas.

---

<sup>21</sup> A la suite de Brest, beaucoup de collectivités ont ainsi ouvert des sites où des habitants racontent leur quartier, d'aujourd'hui et d'hier.

## LA BIBLIOTHEQUE MUNICIPALE DE TOULOUSE SUR FLICKR<sup>22</sup>

Depuis 2008, la bibliothèque municipale de Toulouse diffuse sur le site de partage de photos Flickr le fonds photographique d'un photographe local du 19e siècle, Eugène Trutat. L'ambition initiale du projet, et la raison du choix de Flickr (média web2.0 par excellence), étaient de faire participer les internautes à l'enrichissement du fonds : par des apports de connaissances, des contributions, la qualification de certains clichés restés indéterminés, etc. La conservatrice Jocelyne Deschaux témoigne de l'intérêt que suscitent les photos auprès d'internautes français ou étrangers, et d'une notoriété renforcée : "Flickr nous a aidés à trouver un nouveau public".

Les photos en basse définition sont utilisables sans restriction, tandis que l'usage des photos en haute définition se fait selon la grille tarifaire de la bibliothèque.



### ► Peut-on gagner de l'argent avec les données culturelles du territoire ?

Les fonds d'archives peuvent coûter cher à numériser. Si leur originalité, leur valeur culturelle ou historique, apparaissent importantes, la tentation de les commercialiser pour dégager des ressources financières sera grande. La loi l'autorise dès lors qu'un travail spécifique est réalisé pour les numériser et les rassembler dans une base de données. Mais il n'est sans doute pas inutile de s'interroger sur la réalité des revenus que l'on peut espérer, au regard des autres bénéfices que l'on peut en attendre (attractivité, visibilité du territoire, vie culturelle...) : le "marché" est-il suffisant ? Les agents publics sauront-ils s'y adresser ? Ne coûtera-t-il pas aussi cher que cela rapport de facturer et recouvrer des sommes généralement petites ?

#### 1.2.5. Une contribution au développement économique

Que ce soit directement (au travers de la création de services fondés – en tout ou partie – sur des données publiques) – ou indirectement (au travers des effets positifs de l'ouverture des données sur le dynamisme du territoire), on attend de l'ouverture des données des effets sur le développement économique. Sans clairement fournir sa source, le gouvernement britannique l'estime à plusieurs dixièmes de points de croissance.

Il existe cependant deux manières très différentes de mesurer la contribution des données publiques "ouvertes" au développement économique. Ces deux approches ont des conséquences

<sup>22</sup> Cf. le dossier de la Gazette des communes : <http://infos.lagazettedescommunes.com/48352/I%20E2%80%99exception-culturelle-se-glisse-dans-le-reseau-numerique/>

différentes sur la manière dont les acteurs publics peuvent envisager de mettre leurs informations à disposition.

#### ► **La manière directe : mesurer le "marché de l'informatique publique"**

Dans un premier temps, on peut mesurer le marché de l'information publique en additionnant les chiffres d'affaires des organismes publics (Insee, IGN, Inpi...) et des entreprises privées qui vendent des bases de données juridiques ou économiques, des informations financières, des services cartographiques, etc. En 2007, le GFII estimait le chiffre d'affaires de l'information professionnelle sous forme électronique à 1,6 milliard d'euros, dont l'information publique représentait 60% (950 millions d'euros). Au niveau européen, la Commission européenne estime le marché des données publiques à 27 milliards d'euros en 2010.

Ce marché s'organise aujourd'hui autour d'un petit nombre d'entreprises spécialisées, pour la plupart regroupées au sein du Groupement français de l'industrie de l'information (GFII – voir encadré). L'ouverture plus large, et moins coûteuse, d'un ensemble plus divers de données publiques, peut à la fois favoriser le développement de ces entreprises et permettre à de nouvelles entreprises d'entrer sur ce marché.

Encore faut-il que les conditions dans lesquelles les données sont mises à disposition permettent ce développement. Ainsi, l'Office of Fair Trading, équivalent britannique du Conseil de la concurrence, a publié en décembre 2006 un rapport intitulé *The commercial use of public information*<sup>23</sup>. Ce rapport évalue à 870 millions d'euros les revenus issus en 2005 de l'information publique, soit par vente directe, soit par vente issue de la réutilisation de cette information. Il évalue aussi le manque à gagner, c'est-à-dire le chiffre d'affaires qui serait réalisé si étaient levés les différents freins actuellement mis par les administrations et les organismes publics à la réutilisation de leurs données : prix de vente trop élevé (30 M €), restrictions d'accès (206 M €) et non-accessibilité de l'information (529 M €). Le déficit de l'activité économique est donc évalué à 765 M €, soit 88 % du marché mesurable<sup>23</sup> !

#### ► **La mesure indirecte : l'information publique, matière première de la production de services**

Mais le potentiel économique des données publiques est sans doute beaucoup plus important. On sait que l'économie du numérique ne se résume pas aux entreprises spécialisées, fournisseurs de technologies, opérateurs télécoms, producteurs de contenus ou de services en ligne : toutes les entreprises sont concernées, que ce soit dans les services (services à distance, création de nouveaux services, accès à de nouveaux marchés...) ou dans l'industrie (gains de productivité, optimisation des chaînes d'approvisionnement et de la logistique, personnalisation, accélération des cycles d'innovation...). Il en va de même pour les données publiques : celles-ci peuvent faciliter l'émergence ou la croissance d'entreprise dont elles ne sont pas le produit, mais la matière première : des opérateurs de mobilité pourront proposer des services multimodaux, des acteurs des services urbains pourront mieux organiser leur présence et leur offre, des agents immobiliers fourniront des informations détaillées sur le quartier dans lequel ils proposent des biens, des entreprises réalisent des études d'implantation ou des analyses marketing beaucoup plus précises, etc.

Parce qu'elle touche un périmètre beaucoup plus large, on peut estimer que la contribution indirecte des données publiques à la croissance est bien supérieure au chiffre d'affaires du "marché de l'information publique" décrit plus haut. En revanche, elle se mesure beaucoup plus difficilement, dans la mesure où il n'existe pas de lien direct entre les données et la valeur du service ou du produit final.

#### ► **Deux formes de retour financier pour les acteurs publics**

Pour les acteurs publics, ces deux manières d'estimer la contribution économique des données publiques ont des conséquences différentes. Dans le premier cas, ils peuvent être tentés de tarifer leurs informations en fonction du chiffre d'affaires que les entreprises spécialisées escomptent en tirer<sup>24</sup>. Dans le second cas, la chose est pratiquement impossible et le retour financier pour l'acteur

<sup>23</sup> Source : "La longue marche de l'information publique", Bernard Marx et Ruth Martinez, *Documentaliste – Science de l'information*, 2007 - <http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2007-3-page-218.htm>

<sup>24</sup> Même si la loi invite les acteurs publics à tarifer en fonction du "coût de mise à disposition" et non d'un prix de

public proviendra plutôt des taxes sur l'activité économique des entreprises que de la cession des données. On cherchera alors plutôt à abaisser les barrières économiques, techniques ou pratiques à l'accès aux données afin de favoriser l'activité.

#### **La tarification comme choix économique : l'approche d'un économiste**

L'économiste Olivier Bomsel de Mines-ParisTech décrit trois manières de tarifer l'information publique :

- Comme une rente (royaltie) prélevée sur le bénéfice de l'édition privée : port, octroi, droit de passage...
- Comme une facilité essentielle : au coût moyen incrémental de long terme (le "coût de mise à disposition")
- Comme un bien public offert aux entreprises au même titre que l'environnement institutionnel

Pour lui, le choix tarifaire reflète des priorités économiques : plus il est cher, plus le service public est en monopole, et plus les services apportés par les entreprises privées sont chères pour le consommateur

## Les entreprises, productrices de données publiques ouvertes ?

Le potentiel des données ouvertes est suffisamment important pour que même des entreprises privées, qui n'y sont nullement invitées par la loi, rendent accessibles leurs données de différentes manières dans le but de permettre à d'autres acteurs de créer des services à valeur ajoutée.

Nous avons rencontré page 6 les exemples d'Amazon ou de Google, mais le mouvement peut toucher des entreprises beaucoup plus modestes. La sociologue Valérie Peugeot mentionne l'exemple de la start-up anglaise [Nestoria](#), composée d'une dizaine de personnes et implantée dans 6 pays, dont le service adjoint à la recherche d'annonces immobilières des informations à valeur ajoutée sur le quartier (présence des services, données liées aux transports, équipements scolaires, etc.), récupérées grâce à une collecte de données publiques. En s'alliant à des agences immobilières, Nestoria offre ainsi un service de recherche plus intelligent. Le revenu généré par la consultation des annonces est alors partagé entre les deux entreprises.

Nestoria ne se contente pas de réutiliser des données publiques. Elle recueille également manuellement d'autres données sur les quartiers, et elle met à disposition l'ensemble de ses données, sous la forme d'API ouvertes, afin que d'autres puissent s'en servir et les enrichir. Notons que Nestoria apporte aussi un soutien financier au développement d'OpenStreetMap.

The screenshot shows the Nestoria homepage with a search bar for 'Birmingham'. Below it, a list of 9,483 properties for sale in Birmingham is displayed, with the first six results shown:

| Address  | Price    |
|--|----------|
| King Street, Birmingham                                | £209,950 |
| Upper Marshall Street, Birmingham                      | £140,000 |
| Holliday Street, Birmingham - Lift                     | £187,000 |
| Hugh Road, Birmingham - Reception                      | £139,950 |
| Gladstone Rd, Sparkbrook, Birmingham, W. Midlands, B11 | £144,950 |

Each listing includes a small thumbnail image of the property, its address, price, and a brief description. To the right of the list is a map of Birmingham with several location markers labeled A through F, corresponding to the properties listed. The map also shows various streets, landmarks, and green spaces.

En France, plusieurs entreprises publiques et prouvées dans les domaines des transports, du courrier, de l'environnement, de l'eau, des télécoms..., réfléchissent aux informations qu'elles pourraient mettre à disposition, d'une manière généralement gratuite. Leur objectif n'est pas philanthropique, mais bien économique : elles attendent que l'intervention d'autres entrepreneurs contribuent à améliorer leurs services, à augmenter leur trafic ou à élargir leur zone de chalandise. L'entreprise de transport Keolis a ouvert la voie avec la métropole rennaise, et d'autres suivront.

## **1.3. UN NOUVEL ECOSYSTEME D'ACTEURS INTERDEPENDANTS**

L'ouverture des données publiques ne crée pas un simple face-à-face entre des acteurs publics qui produisent ou détiennent des données, et des acteurs privés qui en tirent profit. Elle crée un véritable "écosystème", dans lequel chaque acteur joue tour à tour plusieurs rôles et contribue au travail des autres.

Ainsi, les administrations sont-elles à la fois fournisseurs de données, et utilisateurs des données publiques produites par d'autres acteurs publics ; les citoyens, les associations et les entreprises ne se contentent pas d'exploiter des données publiques, ils peuvent contribuer à les corriger ou les enrichir. Les chercheurs ou les groupes de citoyens qui produisent par exemple de nouvelles cartographies, ou des analyses inédites des dépenses publiques, produisent des connaissances qui peuvent s'avérer utiles à tous ; à leur tour, les acteurs publics peuvent participer à cette création de connaissance en explicitant le sens de chaque donnée, en expliquant comment elles sont produites. Et ainsi de suite.

### ► **Anciens et nouveaux acteurs**

Certains des acteurs sont bien connus : les institutions publiques – parmi lesquels nous inclurons les entreprises délégataires de services publics –, ainsi que les fournisseurs spécialisés de services d'information professionnelle, qui réutilisent depuis longtemps des données publiques. Mais de nouveaux acteurs joueront également un rôle essentiel dès lors que l'accessibilité des données publiques devient à la fois plus large (touchant de nouveaux types d'informations, provenant de sources plus diverses) et plus aisée.

#### **Les principaux acteurs de l'"écosystème" des données publiques**

|   | Produc-teurs | Réutil-sateurs | Inter-médiaires | Intérêt éco. | Intérêt public |
|---|--------------|----------------|-----------------|--------------|----------------|
| <b>Institutions publiques</b>                 |              |                |                 |              |                |
| Nationales / Continentales ("Gouvernements")  | ++           | +              | +               | +            | ++             |
| Locales                                       | ++           | +              |                 | +            | ++             |
| Opérateurs de services publics                | ++           | +              |                 | ++           | ++             |
| <b>Entreprises</b>                            |              |                |                 |              |                |
| Fournisseurs spécialisés de services en ligne | +            | ++             | ++              | ++           |                |
| Autres entreprises existantes                 | +            | ++             |                 | ++           |                |
| Startups                                      |              | ++             | +               | ++           |                |
| Fournisseurs de technologies                  |              | +              | +               | +            | +              |
| <b>Autres</b>                                 |              |                |                 |              |                |
| Associations, activistes, lobbyistes          |              | ++             | +               |              | ++             |
| Médias, blogueurs                             |              | ++             | +               | +            | ++             |
| Chercheurs                                    | +            | ++             |                 | ++           | ++             |
| Citoyens                                      | +            | +              |                 | +            | ++             |

Il est important de décrire certains de ces acteurs, avec lesquels les institutions publiques n'ont souvent pas l'habitude de travailler, ou en tout cas pas de cette manière.

#### **Startups et entreprises de services**

Le "marché" des données publiques se limitait jusqu'ici pour ainsi dire à des bases de données juridiques, économiques, statistiques et géographiques, exploitées par des acteurs spécialisés : fournisseurs de bases de données, cabinets d'études, systèmes d'information géographiques...

Les nouvelles données mises à disposition intéresseront bien d'autres entreprises. Pensons par exemple aux multiples applications sur les mobiles qui se fondent sur les informations relatives

aux transports publics, à la disponibilité de véhicules ou vélos en libre-service, au trafic ; à l'émergence de "cityguides" qui fournissent également une somme d'information sur la localisation des services et équipements collectifs, leurs conditions d'accès, leurs heures d'ouverture ; ou encore, à l'émergence de services relatifs à l'environnement. Ceux-ci sont souvent fournis par de toutes petites entreprises ; ils s'adressent aux usagers, plutôt qu'à d'autres entreprises ; leur échelle va, selon les cas, du local à l'international ; ils apparaissent, disparaissent, évoluent rapidement ; les responsables de ces entreprises n'ont aucune habitude d'échanger avec les administrations ; des délais excessifs, des exigences administratives lourdes, des coûts d'accès élevés, ou tout simplement une incertitude sur ces trois facteurs, peuvent les tuer. Rien à avoir avec les acteurs professionnels auxquels les administrations s'étaient habituées.

### Fournisseurs de technologies

Les entreprises qui fournissent<sup>25</sup> aux administrations les logiciels ou les services en ligne sur lesquelles se fondent leurs systèmes de production et de gestion vont devoir faire évoluer l'architecture même de leurs produits. Celle-ci devra traiter les données d'une manière beaucoup plus indépendante des traitements qu'elles subiront pour répondre aux besoins des administrations. Il leur faudra prévoir les manières de rendre ces données disponibles à des réutilisateurs, sans bien sûr compromettre l'intégrité des applications propres aux administrations. Dans certains cas, il faudra savoir intégrer des enrichissements qui viennent de l'extérieur ; dans d'autres, faire respecter des conditions d'utilisation. Dans certains cas, on pourra se contenter d'exporter des fichiers et de les rendre téléchargeables ; dans d'autres (grandes bases de données complexes, informations temps réel...), on devra créer des "portes d'entrée" (on parle d'API, ou interfaces de programmation) pour permettre aux réutilisateurs de se "servir" à la source – ce qui obligera parfois à redimensionner les infrastructures informatiques, à prévoir des dispositifs de sécurité, etc.

On peut prévoir que ces exigences conduiront rapidement à faire évoluer les cahiers des charges que produiront les acteurs publics, ainsi que leurs critères de sélection de leurs prestataires.

### "Hackers citoyens"<sup>26</sup> et développeur indépendants

Au-delà des associations citoyennes ou encore des "lobbies" organisés, et souvent bien avant eux, une nouvelle catégorie d'acteurs a joué un rôle décisif dans l'ouverture des données publiques : il s'agit de développeurs indépendants, réunis en groupes souvent informels, militants d'une démocratie plus participative enrichie par les outils numériques. Ils ont dans un premier temps participé à collecter des données et ont incité ainsi les administrations à ouvrir leurs données. Comme l'indique un article publié sur le site Owni.fr, "la société civile a imposé des exemples de bonnes pratiques et à mis sous pression les gouvernements. Ainsi, data.gov.uk et data.gov n'auraient sans doute jamais été créées si des citoyens engagés n'avaient pas déjà développé des applications telles que [TheyWorkForYou.com](#) (au Royaume-Uni) ou [GovTrack.us](#) (aux États-Unis), des cartes des accidents de vélo, ou des sites répertoriant les dépenses publiques."

Les deux portails britanniques et américains n'auraient sans doute jamais été créés "*si des citoyens engagés n'avaient pas déjà développé des applications telles que TheyWorkForYou au Royaume-Uni, ou GovTrack.us aux Etats-Unis, des cartes des accidents de vélo, ou des sites répertoriant les dépenses publiques*"<sup>27</sup>. Aujourd'hui souvent regroupés en associations, ils participent activement à la construction d'outils qui facilitent l'interprétation des données publiques, voire parfois à la collecte de données ou à l'accompagnement des collectivités locales. Cette communauté constitue un réseau informel d'acteurs qui se connaissent entre eux, et où prennent naissance de multiples initiatives, à l'échelle nationale ou même européenne.

En Angleterre l'organisation la plus représentative de ce mouvement est [MySociety](#), regroupant aujourd'hui une dizaine de permanents, et animant à elle seule une dizaine de sites web à finalité démocratique ou de simplification de la vie du citoyen. Cette structure finance ses activités d'intérêt général par des missions de conseil, y compris auprès du gouvernement britannique.

<sup>25</sup> Ou même les associations grâce auxquelles des acteurs publics coproduisent et mutualisent des logiciels, telles que l'Adullact.

<sup>26</sup> Le terme "hacker" a mauvaise presse. Souvent assimilés aux "pirates", les "hackers" sont en fait, historiquement, des techniciens à la fois compétents et animés d'une intention "sociale", capables de modifier un objet ou un mécanisme pour lui faire faire autre chose que ce qui était initialement prévu.

<sup>27</sup> "Les bons ingrédients pour une ouverture des données publiques réussie" : <http://owni.fr/2010/05/31/opendata-12-data-gov-ou-data-gov-uk/>

En France, des associations telles que Regards Citoyens à Paris, LiberTIC à Nantes ou la Quadrature du Net jouent un rôle équivalent, même si leur taille est plus modeste.

### Les associations fédératrices

D'autres associations participent activement à l'ouverture des données publiques, en animant les différents réseaux, en favorisant la rencontre entre les parties prenantes. Elles émanent le plus souvent d'acteurs du numérique, mais à la différence des précédentes, elles ne développent généralement pas elles-mêmes des applications.

Le [Forum Virium](#) à Helsinki, en France le [GFII](#) et la [FING](#), jouent ce rôle de facilitateurs, médiateurs, mutualisateurs : en mettant en contact les différents acteurs entre eux, en permettant que les différents intérêts s'expriment, en partageant expériences et savoir-faire, etc.

### Les "data-journalistes"

La presse a souvent été à l'avant-garde du mouvement d'OpenData : le *Guardian* au Royaume-Uni, le *New York Times* ou le *Los Angeles Times* aux Etats-Unis, ont à la fois milité en faveur de cette ouverture, expérimenté très tôt ce qui pouvait être fait de ces données, et informé sur les initiatives des Etats ou des territoires dans ce domaine.

Ces médias établis, ainsi que d'autres beaucoup plus jeunes tels que le Français [Owni](#), ont donné naissance à une nouvelle forme de journalisme : le "data-journalisme"<sup>28</sup>. Selon Wikipedia, le datajournalisme est "une nouvelle technique journalistique qui consiste à analyser des données complexes (par exemple des statistiques) ou à extraire des informations pertinentes de quantités importantes de données." Il s'agit donc de convertir les données en informations nouvelles, intelligibles, que d'autres modes d'investigation n'auraient souvent pas permis de faire émerger. Pour cela, il faut savoir réunir plusieurs compétences très différentes : journalistes, développeurs informatiques, designers, graphistes, etc.

Quand les données manquent, le datajournaliste peut faire appel au [crowdsourcing](#), c'est-à-dire à la contribution des citoyens. Ainsi les journalistes d'Owni et Le Post ont-ils plusieurs fois lancés des appels à contribution pour réussir, par exemple, à cartographier les villes sous vidéo-surveillance.

**La carte des villes sous vidéo-surveillance, co-produite avec les internautes<sup>29</sup>**



<sup>28</sup> "Journaliste de données", *Internet Actu*, 2010 : <http://www.internetactu.net/2010/07/09/journaliste-de-donnees-data-as-storytelling>

<sup>29</sup> [http://www.lepost.fr/article/2009/12/11/1836005\\_plus-de-2000-communes-sont-videosurveillees-et-la-votre.html](http://www.lepost.fr/article/2009/12/11/1836005_plus-de-2000-communes-sont-videosurveillees-et-la-votre.html)

## ► Un nouveau rôle pour l'acteur public

En ouvrant l'accès à leurs données, et en permettant à d'autres d'en faire des usages imprévus, les acteurs publics découvrent un nouveau rôle : celui qui consiste à la fois à faire vivre un écosystème aussi dynamique que possible (pour que l'ouverture des données soit féconde en connaissances, en services aux citoyens et en vitalité démocratique) et à réguler l'usage de ces données.

Il s'agit en particulier de trouver le bon équilibre entre plusieurs objectifs :

- En ouvrant l'accès à ses données, l'acteur public crée un nouvel espace d'initiative pour les autres acteurs (publics, privés et associatifs). Il doit s'assurer que cet espace est repérable, suffisamment vaste et peu contrôlé pour que ces acteurs l'investissent effectivement. En particulier, il s'agit d'accepter que certains services concurrenceront de fait ceux de la collectivité, que certaines interprétations des données déplairont, etc.
- L'acteur public doit aussi s'assurer que les services essentiels demeurent disponibles pour toute la population. La mise à disposition des données publiques ne signifie pas que les institutions s'en remettent exclusivement à "l'écosystème" pour délivrer ces services, et même d'autres moins essentiels.
- L'ouverture des données peut créer un cercle vertueux grâce auquel des citoyens, des entreprises, des chercheurs, contribueront à l'enrichissement des informations elles-mêmes, ainsi qu'à créer des connaissances et des services nouveaux. Les institutions doivent se mettre à l'écoute de ce que produira le terrain, et l'incorporer dans leurs propres systèmes.
- Mais les acteurs publics ont aussi la responsabilité de s'assurer que la réutilisation des données ne crée par en elle-même des problèmes : données dénaturées, données temps réel diffusées trop tardivement, etc. Certaines données peuvent être accompagnées de conditions de réutilisation, qu'il s'agira de faire respecter. Pour autant, ces conditions doivent être générales et de haut niveau : on ne saurait imaginer une *imprimatur* préalable à toute réutilisation !

Cette "nouvelle gouvernance" aura ainsi un certain nombre de caractéristiques :

## **Perdre un certain monopole public pour gagner en richesse et en inventivité**

Il est coutume de dire que l'ouverture des données publiques marque le passage de l'information publique aux "actifs" publics. On peut se montrer réticent devant la dimension "patrimoniale" du mot, en revanche il est intéressant de considérer que la donnée devient un "objet actif" : c'est-à-dire dynamique, engendrant une chaîne de réactions.

En effet la transparence accrue des données publiques aura inévitablement pour conséquence la perte d'un certain monopole pour l'acteur public : monopole dans la détention de certaines données, dans leur interprétation légitime, ou parfois dans la délivrance de certaines informations et de certains services.

Cette perte de monopole engendrera nécessairement des problèmes. Tout le monde n'exploitera pas les données de manière compétente. Tout le monde n'en tirera pas les conclusions que les élus souhaiteraient. C'est en revanche la condition pour qu'apparaissent de nouvelles voix, pour libérer les capacités d'innovation économique et sociale, pour faire émerger de nouvelles connaissances. Il s'agit au fond d'élargir l'"espace public", qui est à la fois une nécessité démocratique, un carrefour d'échanges économiques, et le lieu dans lequel se déroulent les manifestations...

## **Soutenir, accompagner l'innovation territoriale**

Il s'agit bien pour l'acteur public de mettre des ressources publiques au service de l'innovation sociale et économique, telle qu'elle se crée et s'exprime aujourd'hui. Pour cela, il doit à la fois :

- Stimuler la réutilisation : en l'incitant, la provoquant, en l'organisant, en la publicisant ;
- Accepter de lâcher-prise sur les productions : pour que la créativité des acteurs privés, associatifs, citoyens, et l'innovation s'exprime sans contraintes excessives ;
- Conserver une vue d'ensemble pour s'assurer que l'intérêt général est atteint, que l'accès aux services publics demeure universel, voire en coordonnant des initiatives afin d'en élargir l'accès à un plus grand nombre de citoyens.

L'acteur public est légitime à coordonner et fédérer l'intelligence collective territoriale ; légitime aussi à rappeler les grands principes des services au public, et les besoins et les manques existants.

### **Imaginer des infrastructures sociales et pas seulement techniques**

L'infrastructure technique sous-jacente à l'ouverture des données (bases de données, modes de téléchargement, licences, portails, etc.) est nécessaire à l'écosystème, mais elle ne suffit pas. Elle doit se compléter d'une "infrastructure sociale"<sup>30</sup> : c'est-à-dire les modalités d'accompagnement de la réutilisation, les formes d'animation de la communauté des ré utilisateurs, la facilitation des modes de partage entre réutilisateurs.

Si l'acteur public perd une partie du contrôle de la réutilisation, cela ne veut pas dire que celle-ci se déroule en dehors de tout contexte ou influence politique, socioculturelle, etc. Au contraire, l'analyse des projets et initiatives existantes – par exemple en Angleterre - montre que ceux-ci se développent au sein de réseaux d'acteurs ayant besoin de collaborer entre eux (ne serait-ce qu'en raison de la diversité des compétences requises). Charge à l'acteur public de favoriser la mise en contact des acteurs entre eux, de faire valoir les besoins et les manques de services identifiés, voire d'influencer l'orientation du développement des services (par des concours, des formes d'animation ou de soutien spécifique).

Le concours lancé dès 2008 par le gouvernement britannique, "[Show us a better way](#)" ("Montrez-nous une meilleure voie") est symbolique de cette nouvelle manière de faire. Elle a d'ailleurs inspiré de nombreuses opérations similaires comme "Apps for Democracy" (Washington, Finlande) ou le "Concours de création d'applications" de Rennes?

### **Offrir une reconnaissance aux travailleurs de l'ombre**

Les bonnes idées ou pratiques citoyennes ont besoin de relais auprès des agents et des cadres des institutions publiques. Ces "travailleurs de l'ombre" existent. Aujourd'hui, en France, une poignée d'agents volontaires sont souvent à l'origine des initiatives de leur ministère ou de leur collectivité, alors même que ces chantiers ne sont pas toujours officiellement inscrits dans leurs missions... Ils doivent avancer en même temps que faire la preuve à leur hiérarchie de l'utilité de ces projets.

En Angleterre, ces fonctionnaires de l'ombre ont là aussi joué un rôle central : l'expertise qu'ils avaient acquises a donné confiance "aux leaders politiques dans la réussite de stratégies d'ouvertures des données publiques. Les dirigeants ont ainsi eu les moyens et la motivation pour passer outre l'inertie institutionnelle"<sup>31</sup>.

### **Evaluer la création de valeurs en imaginant de nouveaux indicateurs**

Il reste néanmoins une difficulté de taille à passer : l'évaluation de l'innovation produite grâce à la réutilisation des données publiques. La valeur économique n'est pas le seul indicateur pertinent. Quantifier, identifier le bénéfice global pour la société de telles ou telles innovations requerra des indicateurs plus divers et plus "sociétaux".

L'évaluation est d'autant plus complexe que les retombées se constatent en deux, voire trois temps :

- La création d'activités et de services, commerciaux ou non ;
- L'usage par les habitants et les entreprises, et ses conséquences en termes de vie quotidienne, de productivité, etc. ;
- Le retour pour les pouvoirs publics, parfois mesurable (recettes directes des données publiques, recettes fiscales), parfois beaucoup moins (qualité de vie, environnement, vitalité démocratique, cohésion sociale, compétitivité du territoire...).

<sup>30</sup> Pour reprendre l'expression de Tim Davies dans son excellent rapport "How is open government data being used in practice" : <http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf>

<sup>31</sup> <http://owni.fr/2010/05/31/opendata-12-data-qov-ou-data-qov-uk/>

## **Territoires 2.0 : gérer une plate-forme d'innovation ouverte**

Dans leur ouvrage "La ville 2.0, plate-forme d'innovation ouverte<sup>32</sup>", Daniel Kaplan et Thierry Marcou de la Fing décrivent les 5 fonctions à remplir pour gérer une "plate-forme d'innovation ouverte" dans un territoire. Extraits :

### **Faciliter le partage et l'accès aux ressources partagées**

L'infrastructure d'innovation doit assurer la mise à disposition des ressources partagées (notamment des informations et des services), sous des formes pertinentes pour ceux qui en ont besoin. Une telle tâche peut recouper plusieurs fonctions :

- La gestion d'un annuaire de ressources,
- L'assistance aux détenteurs d'informations, par exemple pour formater ou cataloguer leurs ressources
- L'hébergement de données, ou autres prestations techniques.

### **Réduire les coûts pour tous les acteurs**

En gérant certaines ressources communes, l'infrastructure d'innovation peut réduire les coûts d'investissement et d'exploitation associés aux innovations urbaines. Parmi les fonctions qui pourraient être remplies de cette manière, on peut énumérer :

- La mise à disposition d'infrastructures techniques communes
- La gestion, voire la production de référentiels communs tels que des fonds de cartes, des modèles 3D, des typologies, des vocabulaires, etc.
- La gestion de services communs tels que l'identification et l'authentification des utilisateurs.

### **Faciliter les partenariats**

L'innovation urbaine ouverte étant par nature partenariale, elle aura besoin de facilitateurs.

Des réunions ou des manifestations destinées aux innovateurs complètent généralement ce dispositif. Certaines ont pour objet de familiariser une communauté d'innovateurs avec les ressources mises à leur disposition et de les inciter à se regrouper entre projets similaires. D'autres opéreront la rencontre entre des porteurs de projets et des institutions, des entreprises, des acteurs urbains installés des réseaux associatifs et citoyens...

### **Favoriser la rencontre entre les projets et les utilisateurs**

Les petits innovateurs rencontrent souvent les plus grandes difficultés pour associer des utilisateurs à leurs projets, puis pour accéder aux publics qu'ils visent, en phase d'expérimentation ou de lancement. Les infrastructures d'innovation ouverte peuvent intervenir à ces deux étapes.

Pendant la période de conception, elles peuvent favoriser la rencontre entre des innovateurs et des communautés d'utilisateurs actifs, dans des démarches de co-conception.

Pendant la période d'expérimentation, l'infrastructure d'innovation peut assurer une fonction d'interface entre l'innovateur et les utilisateurs. Certaines municipalités envisagent ainsi de recruter elles-mêmes des testeurs pour certains prototypes d'innovations urbaines qu'elles jugent intéressants.

L'infrastructure d'innovation peut ensuite, soit fournir un espace de visibilité aux projets innovants (sur un portail, par exemple, ou encore *via* des espaces d'affichage dans la ville), soit ouvrir les portes des grands médias et portails de services actifs dans la ville. D'autres systèmes, tels que des labels, des trophées et autres manifestations publiques, peuvent aussi permettre aux projets les plus intéressants d'attirer l'attention des médias et des grands acteurs.

### **Réguler les interactions entre les acteurs**

Enfin, l'infrastructure d'innovation peut jouer un rôle important dans la régulation des relations entre les acteurs de l'écosystème d'innovation urbaine.

<sup>32</sup> Fyp Editions, 2009, téléchargeable ici : <http://fing.org/?La-Ville-2-0-plateforme-d-innovation-ouverte>

Elle est d'une part un lieu permanent de rencontre et d'interaction entre ces acteurs, ce qui peut contribuer à un climat de confiance et aider à résoudre d'éventuels problèmes suffisamment tôt et d'une manière informelle.

Elle contribue par ailleurs de plusieurs manières à la régulation des échanges. Elle informe sur les conditions (y compris économiques) d'utilisation des ressources partagées. Elle participe à la capitalisation des retours d'expérience.

Enfin, l'infrastructure d'innovation constitue un formidable lieu d'observation de la dynamique d'innovation territoriale, ainsi que des pratiques et des usages dont elle suscite l'émergence. Ce rôle doit lui aussi être pris au sérieux et aboutir à un retour d'information de qualité au bénéfice de tous les participants.

## 1.4. UN CADRE JURIDIQUE TRES FAVORABLE

### Un ouvrage complet, le guide Juridique d'AEC

Nous ne rentrerons pas dans ce présent Guide pratique dans tous les détails de la loi et renvoyons, pour ceux qui désirent approfondir, au Guide juridique d'AEC, *Les données publiques*<sup>33</sup>. Ce guide est librement téléchargeable.

#### L'objectif du présent guide juridique est double :

- *Rappeler de façon synthétique et pratique les obligations légales auxquelles les acteurs publics sont soumis en matière d'accès et de réutilisation des données qu'ils produisent et détiennent dans le cadre de leurs missions de service public (les données publiques au sens propre).*
- *Accompagner ces acteurs publics dans leur mise en conformité avec les prescriptions légales.*
- *Enfin, ce guide s'adresse aux acteurs publics désireux de s'engager dans une démarche volontariste d'ouverture de leurs données, en leur proposant un cadre juridique et pratique afin d'opérer cette libération en toute sécurité.*



La législation française encadre clairement la réutilisation des données publique à travers le chapitre II de loi 78-753 de 1978, dite loi CADA. Nous retenons tout d'abord deux choses essentielles de ce texte :

- sauf exceptions – vues précédemment –, toute donnée collectée dans le cadre d'une mission de service public est réutilisable
- la loi française introduit un **droit opposable de réutilisation des données publiques** : par défaut, réutiliser une information publique est un droit, que n'importe qui peut faire valoir auprès de n'importe quel acteur public

Au delà, la loi de 1978 exprime des droits et des obligations pour l'acteur public comme pour le réutilisateur.

#### ► Les droits et obligations de l'acteur public

- Les obligations :
  - Mettre à disposition les données publiques collectées dans le cadre de ses missions – sauf exceptions prévues par la loi
  - Constituer un répertoire des informations publiques, en ligne s'il dispose d'un site web
  - Ne pas consentir d'accords exclusifs de réutilisation des données publiques
  - Rédiger ou utiliser une licence type lorsque des données sont soumises à redevance
  - Si l'acteur public choisit de demander une redevance sur ses données, il a l'obligation d'en publier les modes de calcul et de la rendre "non discriminatoire" : elle doit être identique pour deux acteurs exploitant les mêmes données dans les mêmes conditions
- Les droits :
  - Percevoir une redevance pour la mise à disposition des données publiques ; cette redevance doit cependant être proportionnée au coût de diffusion
  - Faire respecter les conditions de la licence ou les obligations légales liées à la réutilisation
  - Choisir le format, la méthode et la fréquence de mise à disposition des données
  - l'acteur public peut apporter des restrictions à la réutilisation, "toute restriction devant impérativement être motivée par l'intérêt général, respecter le droit de la concurrence et être proportionnée à la réutilisation envisagée"

<sup>33</sup> [www.aecom.org/Vous-informer/Juridique-TIC/Guides-juridiques/Les-donnees-publiques-decembre-2010](http://www.aecom.org/Vous-informer/Juridique-TIC/Guides-juridiques/Les-donnees-publiques-decembre-2010)

## ► **Les droits et obligations des réutilisateurs**

- Les droits :
  - Le réutilisateur bénéficie d'un droit personnel et non exclusif de réutilisation des données publiques
  - *Sans mention de redevance*, la réutilisation des informations publiques est gratuite
  - *Sauf mention contraire*, la réutilisation des informations est libre, sans autre restriction que celles prévues par la loi
- Les obligations :
  - Indiquer la source des données et leur date de dernière mise à jour<sup>34</sup>
  - Ne pas dénaturer le sens des données<sup>35</sup>
  - Ne pas réutiliser d'informations à caractère personnel, sauf autorisation explicite des personnes concernées

On le voit, le législateur français a choisi de promouvoir sans ambiguïté la réutilisation des données publiques.

---

<sup>34</sup> Sauf permission explicite et motivée, actuellement du domaine de l'exception.  
<sup>35</sup> Idem note précédente.

## **2. COMMENT S'Y PRENDRE ?**

### **2.1. ENGAGER UNE POLITIQUE DE REUTILISATION DES DONNEES PUBLIQUES**

#### **2.1.1. Une démarche à la fois volontariste, pragmatique et ouverte**

La réutilisation des données est un droit qui s'impose aux acteurs publics, mais ceux-ci ont tout intérêt à engager une démarche volontariste, sans attendre. Il s'agit en effet de maîtriser le cheminement plutôt que le subir : identifier les priorités plutôt que de se les faire dicter ; expérimenter pour prendre les bonnes décisions au lieu de livrer sous la contrainte ; industrialiser la production des données plutôt que risquer de se voir dépassé par la demande.

La démarche gagne à être pragmatique. Tout d'abord, il s'agit d'un sujet nouveau et prospectif. On aura donc tout intérêt à avancer progressivement. De petits réajustements valent mieux que de grands changements de cap. L'acteur public a tout intérêt à démarrer dans un mode expérimental, avec des cycles courts, qui permettent notamment de prendre le temps de pérenniser ce qui marche et de réfléchir à ce qui ne marche pas. Une démarche pragmatique permet ensuite de réduire un champ qui est potentiellement immense, en ne traitant au départ que ce qui offre les meilleures chances de résultats. Elle démontre la volonté de l'acteur public et permet d'anticiper des logiques de confrontations : tout le monde peut admettre qu'une institution qui s'engage sans réticence ne soit pas non plus en mesure de tout mettre à disposition tout de suite. Le pragmatisme est la base de tous les exemples réussis actuellement, notamment dans les pays anglo-saxons.

Dernier point, mais non des moindres, l'ouverture des données publiques relève d'un processus d'innovation ouverte : il s'agit de donner libre cours à l'imagination des acteurs économiques et sociaux, pour qu'ils créent de nouveaux services, de nouvelles connaissances, de nouvelles interprétations. Dans ce contexte, la diffusion des données publiques ne fait que la moitié du chemin et c'est seulement la réutilisation de ces données qui crée de la valeur. L'acteur public ne peut pas, légalement, choisir lui-même qui pourra réutiliser quelles données, à quelles fins. En revanche, il est important, d'un côté, de stimuler la demande de données pour que l'ouverture produise effectivement des résultats ; et de l'autre, d'orienter autant que faire se peut les réutilisateurs vers des usages socialement et économiquement féconds. Le dialogue et la collaboration entre les acteurs publics et les réutilisateurs sont donc des conditions premières de réussite.

Ces processus d'innovation ouverte, s'ils sont bien connus dans l'industrie<sup>36</sup> – notamment dans celle du numérique – sont moins bien connus dans les administrations, quand ils ne sont pas même tout simplement vécus comme dérangeants. Pourtant, dans pareil processus il n'y a pas d'autre option que de savoir lâcher prise, d'admettre que les résultats ne soient pas donnés à l'avance, d'accepter d'être surpris ! Cela ne veut pas dire que l'acteur public doive se résigner à ne rien maîtriser du tout : il a un vrai rôle dans l'organisation de "l'écosystème" qui s'organise autour des données publiques<sup>37</sup>.

<sup>36</sup> Voir par exemple <http://www.internetactu.net/2010/09/23/eric-von-hippel-il-y-a-2-a-3-fois-plus-dinnovations-de-la-part-des-consommateurs-qu'il-ny-en-a-dans-lindustrie/> et <http://www.internetactu.net/2009/06/26/les-innovations-ouvertes-sont-elles-compatibles-avec-les-systemes-dinformation/>

<sup>37</sup> Voir à ce titre <http://www.internetactu.net/2010/08/30/concevoir-pour-perdre-le-controle/> ainsi que l'ouvrage : <http://www.internetactu.net/2010/06/01/pour-un-design-de-politiques-publiques/>

## 2.1.2. Les grandes étapes de la démarche

Selon la nature de l'acteur public (Etat, collectivité locale, établissement public...), son activité, ses moyens, son système d'information, l'ouverture des données publiques prendra des formes très différentes. On peut cependant, pour tous, distinguer trois grandes étapes dans cette démarche de création de valeur à travers la réutilisation des données publiques.

### ► Etape 1 : Esquisser les enjeux, repérer les potentiels de création de valeur

La première étape s'attache à un travail de compréhension du sujet, d'écoute et de dialogue avec les réutilisateurs potentiels, notamment pour évaluer avec eux les potentiels de création de valeur. Elle consiste en un repérage fin des acteurs clé et, en particulier, des premiers réutilisateurs potentiels : qui sont ces gens qui pourraient vouloir de vos données ?

Une démarche simple peut consister à faire le tour des "têtes de réseaux" de votre territoire ou de votre domaine d'action afin de comprendre les acteurs et les problématiques qui pourraient être impactés par la diffusion des données publiques. Ne vous en tenez pas aux acteurs les plus institutionnels, qui ne sont pas toujours les plus innovants : pensez aux réseaux d'entrepreneurs, aux chercheurs, aux associations dynamiques, aux communautés de développeurs informatiques "citoyens", aux artistes...

Ainsi repérés, un bon moyen pour démarrer le dialogue avec les réutilisateurs potentiels consiste à effectuer avec eux le recensement des données dignes d'intérêt. Ce recensement est guidé par le pragmatisme, on regardera plus spécifiquement, au moins pour un début, aux données qui déclenchent des réflexes créatifs : "Ah si j'avais ces données là, je pourrais...". Il permet également, toujours avec les réutilisateurs potentiels, de prioriser ce qui pourra être ouvert, en tenant compte :

- des contraintes de l'acteur public : d'ordre technique, financier, juridique, etc.,
- des motivations et des contraintes des réutilisateurs potentiels,
- et enfin, des autres sources possibles de données ouvertes (autres sources publiques, données privées, données coproduites par les utilisateurs...).

Cette étape conduit aussi à s'interroger sur les partenariats utiles à envisager pour la suite des opérations. Dans le cas de données de transports, par exemple, une administration unique est rarement la seule détentrice de toutes les données utiles. Des partenariats croisés avec d'autres administrations, avec des opérateurs privés, peuvent donc être précieux, si ce n'est indispensables !

### ► Etape 2 : Ouvrir des premiers jeux de données

La seconde étape consiste à ouvrir des données, à livrer progressivement mais régulièrement et aussi à le faire savoir. Nous détaillons cette étape, relativement "technique", dans la partie 2.3, p. 47.

Le premier "lâcher de données<sup>38</sup>" peut être limité et simplifié, ce qui ne signifie pas que les conditions de base que nous décrivons plus bas (documentation des données, formats ouverts ou au moins communs, licences ou conditions d'usage claires et simples) ne s'applique pas : si le but est d'amorcer une dynamique, il vaut mieux abaisser les barrières...

Il est plus utile de s'intéresser à un petit nombre de données en profondeur plutôt qu'à un nombre très large, mais en surface. Traiter à fond un petit nombre de données permet de se confronter à toutes les problématiques juridiques et techniques de la publication avant d'appliquer l'expérience acquise à d'autres jeux de données. Il ne s'agit pas pour non plus de passer plusieurs mois à vouloir documenter chaque information, à produire la meilleure API possible, etc. Mêmes si elles ne sont pas *parfaites*, des livraisons progressives mais régulières assurent un *feedback* régulier des réutilisateurs et permettent de corriger rapidement des problèmes qui coûteront autrement plus cher s'ils sont détectés en fin de process.

---

<sup>38</sup> Pour reprendre la jolie expression inventée par l'association Regards Citoyens

### ► Etape 3 : Accompagner : stimuler, valoriser, aider, dialoguer

La dernière étape est au moins aussi importante que les précédentes. Ouvrir des jeux de données ne suffit pas : sans un minimum d'accompagnement, cela ne produit pas grand chose<sup>39</sup> et l'acteur public minimise ses chances de voir émerger des usages réellement innovants de ses données. Il doit également tâcher d'orienter, de stimuler et de valoriser les usages de ses données.

Les moyens sont nombreux :

- Certains sont classiques et bien connus des acteurs publics comme les appels à projets, l'animation de réseaux d'innovateurs, les rencontres publiques...
- D'autres formes, plus volontaristes ou plus innovantes, peuvent être plus pertinentes. Les **concours** ont été expérimentés aux États-Unis, en Grande-Bretagne, au Canada et en Australie – ainsi qu'à Rennes : les résultats sont souvent très intéressants, comme à Ottawa où, lancé en septembre 2010, Apps4Ottawa<sup>40</sup>, a vu 101 applications<sup>41</sup> concourir pour 50 000\$ canadiens de prix !
- Les **ateliers créatifs** sous toutes leurs formes sont également un moyen très efficace : "Barcamps", forums créatifs, "coding parties", compétitions amicales, etc. Si l'acteur public n'est pas toujours très au fait de ces méthodes, les réseaux d'innovateurs ou des médiateurs spécialisés les connaissent bien. Ces ateliers créatifs sont particulièrement utiles pour sortir des usages trop classiques et pour faire émerger des projets particulièrement innovants.

Plus en aval, la valorisation des projets existants peut prendre diverses formes. Au delà de la labellisation d'applications de qualité, démarche bien connue mais un peu lourde à mettre en œuvre, elle peut prendre la forme d'une **galerie<sup>42</sup> des usages et/ou des applications**. Idéalement publiée à proximité de la description de la démarche et des jeux de données ouverts, elle permet de comprendre concrètement tout le potentiel créatif des données et montre synthétiquement les résultats de l'opération. Cette galerie peut aussi mettre en lumière des réalisations particulièrement emblématiques, choisies par l'acteur public, par un jury ou par un vote du public.

L'acteur public portera une attention toute particulière à la mise en relation entre certains réutilisateurs imaginatifs mais dépourvus de moyens de compétences techniques, et les acteurs capables de réaliser (ingénieurs, start-ups, "geeks"...), permettant ainsi à des problématiques de rencontrer des solutions insoupçonnées.

<sup>39</sup> "Ils ont cru que la communauté allait se bâtir toute seule" se plaint un critique des expériences (médiocres) d'Edmonton, Toronto et Vancouver : <http://weait.com/content/tragedy-edmonton-vancouver-open-data>

<sup>40</sup> [www.apps4ottawa.ca/](http://www.apps4ottawa.ca/)

<sup>41</sup> [www.ottawacitizen.com/technology/City+Ottawa+contest+attracts+double+expected+entries/4065650/story.html](http://www.ottawacitizen.com/technology/City+Ottawa+contest+attracts+double+expected+entries/4065650/story.html)

<sup>42</sup> Ou "showcase" comme disent les anglo-saxons. On verra par exemple celui de San Francisco : <http://datasf.gov/showcase/>

## **2.2. DEFINIR SA POLITIQUE D'OUVERTURE ET DE DIFFUSION DES INFORMATIONS PUBLIQUES : LES GRANDS CHOIX A FAIRE**

Quelles données publier ? Sous quelles conditions et dans quelles limites ? Gratuites ou payantes ? Certaines de ces questions sont bien balisées comme celle, en amont, de la possibilité même de réutiliser des données publiques. D'autres questions, comme par exemple la problématique de la redevance, font l'objet de discussions parfois vives. Cette partie vise à éclairer les choix des collectivités territoriales, sans pour autant les faire à leur place, dans trois domaines : la nature des données à ouvrir, les conditions juridiques de leur mise en accès, et l'éventuelle tarification de ces données.

### **2.2.1. Quelles données ouvrir ?**

#### ► **Ce que dit la loi**

Du point de vue de la loi, on le rappelle, toutes les données produites ou reçues par un acteur public dans l'exercice de ses missions de service public ont vocation à devenir réutilisables, à l'exception :

- des données personnelles,
- et des données susceptibles de porter atteinte à la sécurité publique.

Sont également exclues de l'application "automatique" de la loi, en matière de réutilisation :

- Les données auxquelles sont associées des droits d'auteur, sauf bien sûr si cette réutilisation est explicitement consentie,
- Les services publics industriels et commerciaux, comme par exemple les transports. Cela n'interdit pas aux acteurs publics de *choisir* de rendre ces données réutilisables, ou de l'imposer dans leurs cahiers des charges aux entreprises délégataires de ces services.

#### **Les panneaux de signalisation, une donnée sensible ???**

Qui penserait trouver par exemple des données sensibles dans les panneaux de signalisation urbains ? L'un d'entre eux, pourtant, signale les zones d'arrêt des convois de fonds. Cette information, même si elle est publique, peut être sensible.



#### ► **Quelles données méritent-elles d'être ouvertes ?**

Il est sans doute bien rare qu'une donnée n'ait pas quelque intérêt à être ouverte. L'expérience des pays anglo-saxons et les témoignages des réutilisateurs nous permettent d'affirmer que quels que soient sa forme, toute donnée ou presque peut faire l'objet d'une réutilisation pertinente. C'est pourquoi il est particulièrement utile d'associer très tôt les réutilisateurs potentiels au recensement des données. Les données les plus pertinentes ne sont pas toujours celle que l'on imagine au départ !

Dans ce cas, comment choisir quelles données ouvrir en priorité ? Plusieurs critères de choix peuvent conduire la réflexion :

- **Les priorités publiques** : la mise à disposition de données peut contribuer à mobiliser les énergies pour atteindre un but collectif. Par exemple, tel acteur public pourrait vouloir sélectionner les données qui peuvent avoir un impact concret sur l'accessibilité de la ville et de services aux personnes à mobilité réduite : cartographie des services disposant de dispositif d'accessibilité, etc.
- **La volonté de soutenir une filière économique** : en ouvrant ces données, la collectivité peut fournir un terrain d'expérimentation à des entreprises d'un secteur qu'elle souhaite soutenir, ou pour lequel elle souhaite devenir attractive.
- **La demande** : si la demande spontanée aujourd'hui assez faible, la situation pourrait rapidement changer. Un nombre croissant d'acteurs associatifs et économiques est informé de ce qu'il se passe en France et dans le monde autour des données publiques. Des acteurs

militants s'en saisissent, des entreprises les réclament pour développer leurs services – par exemple dans des domaines tels que la mobilité, ou les services géolocalisés. Et localement, en interrogeant les acteurs, on fera aisément émerger des attentes.

- **Le pragmatisme**, enfin : certaines données sont plus faciles à publier que d'autres (elles sont déjà informatisées d'une manière qui en facilite la mise à disposition) – on aura intérêt à travailler avec les services informatiques du territoire, pour identifier non ce qui difficile (les informaticiens savent très bien faire ça !), mais au contraire ce qui est facile. D'autres données pourraient rapidement susciter l'émergence d'applications, voire bénéficier d'un retour des utilisateurs qui les enrichiront. Dans d'autres cas, enfin, des opérateurs extérieurs, liés ou non à la collectivité, peuvent se montrer volontaires : à Rennes, c'est au départ l'opérateur des vélos en libre-service qui a ouvert ses données, avec le soutien de la métropole. L'acteur public a ensuite étendu la démarche à un grand nombre d'autres données.

### ► Des données intuitivement plus importantes que d'autres ?

Le champ des données publiques réutilisables est donc immense. Il est encore difficile d'en dresser une liste approchant de l'exhaustivité.

#### **Que peut-on partager dans une ville ? (liste non exhaustive)**

| <b>Des informations et des données</b>   |
|--|
| <ul style="list-style-type: none"> <li>• La description du territoire (cartes, cadastre...)</li> <li>• Des fonds documentaires (études, réglementation, statistiques, archives...)</li> <li>• Les données de la décision publique (projets, enquêtes, délibérations, subventions...)</li> <li>• Le fonctionnement des réseaux urbains (eau, énergie, transports, logistique, télécoms...)</li> <li>• La localisation et les horaires d'ouverture des services et des commerces</li> </ul>  |
| <b>Des applications et des services</b>  |
| <ul style="list-style-type: none"> <li>• Des systèmes d'information géographiques (SIG)</li> <li>• Des modèles (représentation, prévision...)</li> <li>• Des applications permettant de calculer des droits (sociaux, etc.)</li> <li>• L'occupation des ressources et des capacités (voirie, bâtiments, espaces, parkings...)</li> <li>• Des mesures (environnement, trafic...)</li> <li>• Des événements (culture, sports, citoyenneté...)</li> <li>• Des informations touristiques, culturelles,</li> <li>• Les flux urbains (circulation...)</li> <li>• Des données de surveillance</li> <li>• ...</li> </ul> |

On peut cependant penser que certaines données peuvent revêtir une importance particulière. A partir de l'expérience d'autres collectivités françaises et européennes, on en proposera ici une première liste :

- Les codes et les règles produits par l'administration : ces informations constituent le socle de l'information administrative et la base de nombreux usages. Nous pensons par exemple : aux codes des communes, aux codes identifiant donnés par l'administration pour normer, ou identifier des produits des entités économiques etc. ; aux codes et ou normes propres à certains produits de consommation, de santé publique (médicaments, produits chimiques, etc.) ; aux codes Siren, Siret, etc.
- Les données devenues essentielles pour des domaines transverses : par exemple, les points de l'IGN, utilisés chaque jours par des milliers de géomètres pour établir des mesures correctes ; ou encore, la localisation des services publics.
- Les données de base de la citoyenneté et de la vie politique : limites géographiques des circonscriptions administratives, des circonscriptions politiques ; noms des représentants politiques et leur territoire ou leur domaine d'action ; bureaux de vote...
- Les données nécessaires au débat démocratique: budgets, marchés publics, délibérations, études et statistiques...
- Les données utiles à une meilleure gestion des énergies et des ressources : pics de consommation, localisation des équipement de recyclage des services de retraitement des déchets et leurs horaires d'ouverture ;

- Les données utiles à la santé et la sécurité publique : données épidémiologiques, pollution, adresses des services de santé, etc.
- Les données culturelles et touristiques, etc.

La réutilisation de ces données devrait être aussi aisée et libre que possible.

### ► **Les mauvaises raisons de ne pas mettre à disposition des données**

En dehors des cas prévus par la loi, les raisons apparentes de ne pas partager sont nombreuses. Là encore, l'expérience sur le terrain conduit à penser qu'elles sont souvent mal fondées, voire qu'elles servent de prétexte à une réticence de fond à partager l'information

#### Mes données ne sont pas de bonne qualité, elles comportent des erreurs

Qui prétend ne jamais en faire ? Et même avec de nombreuses erreurs, les données peuvent tout de même intéresser les réutilisateurs : certaines sont très faciles à corriger, elles peuvent intéresser à titre de comparaison, etc. En outre, on constate que la publication de données peut, dans certains cas, conduire les utilisateurs eux-mêmes à les corriger, au plus grand bénéfice de l'acteur public. Le producteur aurait en revanche intérêt, pour tout jeu de données, à indiquer un taux d'erreur, même grossier.

#### Mes données ne sont pas de bonne qualité car elles sont incomplètes.

Il est très probable qu'elles soient en réalité complètes sous un certain angle : par exemple restreintes à une zone géographique bien définie. Mêmes "incomplètes", ces données peuvent être utiles aux réutilisateurs pour initier un travail. Là encore, le producteur doit documenter clairement la couverture de ses données.

#### Mes données ne sont pas de bonne qualité car elles ne sont pas assez précises.

La précision recherchée ne dépend que de l'usage. Aux réutilisateurs de dire si la précision leur convient.

#### Ces données sont de bonne qualité mais trop anciennes.

Des données périmées pour leur usage initial peuvent néanmoins intéresser de nombreux acteurs – chercheurs, étudiants, consultants... – désireux d'effectuer par exemple des travaux de synthèse, d'analyse historique, etc.

#### Ces données sont de bonne qualité mais leur mise à jour n'est pas assez régulière.

Là encore, elles peuvent présenter un intérêt pour des usages d'analyse comparative, de travaux de synthèses ou autres.

#### Le format de ces données n'est pas standard.

Si le format est documenté, ce n'est pas un problème. Si les réutilisateurs y voient un intérêt, convertir des données bien documentées d'un format à un autre n'est pas si difficile.

#### Mais le format de ces données est mal documenté !

Un réutilisateur vraiment intéressé trouvera toujours moyen de le faire a posteriori. Peut-être une bonne occasion pour vous de dialoguer avec ces réutilisateurs pour produire une documentation de qualité.

#### Ces données n'intéressent personne !

Chiche ! Dans tous les cas, ne pas préjuger à l'avance qu'elles n'intéressent personne.

## 2.2.2. Les conditions juridiques de réutilisation des données

La loi encadre la réutilisation des données publiques en précisant les droits et obligations des acteurs publics comme des réutilisateurs (voir chapitre 1.4, p. 34). Pour autant, les producteurs de données peuvent vouloir, à l'intérieur de ce cadre légal, effectuer des choix plus précis. Tel voudra rendre aussi libre que possible la réutilisation ; tel autre voudra imposer des conditions strictes en matière d'intégrité des données, d'identification de la source, etc. Tel autre choisira de différencier les conditions qui s'appliquent aux usages commerciaux et non-commerciaux.

Préciser les conditions d'usage de ces données requiert des compétences juridiques. Dont toutes les collectivités ne disposent pas aisément. C'est pourquoi, afin de faciliter leur choix, différentes organisations ont élaboré des **cadres juridiques "clés en mains"** qui ont pris la forme, soit de licences, soit de conditions générales d'utilisation<sup>43</sup>.

Plusieurs points peuvent en particulier faire l'objet de choix spécifiques, qui conditionneront le choix de tel ou tel cadre juridique :

- Les restrictions à la réutilisation, "toute restriction devant impérativement être motivée par l'intérêt général, respecter le droit de la concurrence et être proportionnée à la réutilisation envisagée"<sup>44</sup>
- Le choix de percevoir une redevance pour la réutilisation des données
- Le choix d'ouvrir plus largement certaines possibilités, qui doivent être explicitement indiquées : autoriser l'altération des données, lever l'obligation d'indiquer la source des données leur date de mise à jour...
- La volonté de s'inscrire dans un cadre juridique commun avec d'autres acteurs, par exemple dans un domaine professionnel (on pense aux données de mobilité), ou en choisissant une licence de portée internationale.

### ► Pourquoi c'est important

Un cadre juridique n'est pas seulement là pour *protéger* mais aussi pour *faciliter la rencontre entre le détendeur des données et ses réutilisateurs*. Ainsi, un cadre trop restrictif peut conduire à une valorisation médiocre voire, nulle. Un cadre trop imprécis peut créer de l'incertitude, ou favoriser des usages néfastes qui nuisent à tous les acteurs – pensons à l'impact d'informations sur le trafic, si elles sont rediffusées trop tardivement. Un cadre très précis mais incompréhensible par des non-spécialistes chassera les petits acteurs, souvent les plus innovants...

Il est généralement utile de choisir un cadre juridique connu (voir ci-dessous) plutôt que de réinventer la roue. Produire soi-même ses conditions d'utilisation ou sa propre licence est contre-productif. Un cadre juridique particulier est un objet juridique non identifié, obligeant tous les utilisateurs à l'analyser et le discuter, faisant ainsi perdre un temps précieux à tout le monde. Vous avez en outre toutes les chances de commettre des erreurs alors qu'il existe de nombreux cadres déjà éprouvés

On le voit, **le choix d'un cadre juridique a donc un impact réel sur le succès des usages de données publiques.**

Mais comment choisir ? Nous nous proposons de faire le tour des différents critères et fonctionnalités qui distinguent les cadres juridiques "prêts à l'emploi".

---

<sup>43</sup> Pourquoi parlons-nous de "cadre juridique" et pas de "licence de réutilisation" ? Tout simplement parce qu'une licence n'est pas le seul cadre juridique possible. La réutilisation n'a pas nécessairement besoin de licence, la loi de 1978 fournit un cadre par défaut. Plus récemment, l'APIE a produit des "conditions générales d'utilisation" qui jouent un rôle proche de la notion de licence.

<sup>44</sup> Les données publiques. Guide juridique et pratique, AEC.

## ► Les principaux cadres juridiques en France

Nous avons retenu 8 cadres juridiques, ou plutôt 7 cadres et une "situation par défaut".

**La situation par défaut** est celle qui voit l'acteur public ne mentionner aucun cadre : c'est alors la loi du 17 juillet 1978, modifiée en 2005, qui s'applique<sup>45</sup>. Si l'acteur public fait le choix de s'en tenir à la loi, il n'est pas indispensable de le préciser, mais cela facilitera néanmoins la vie des réutilisateurs. A noter que la décision d'appliquer une redevance ou des conditions restrictives de réutilisation impose la production d'une licence.

D'autres cadres juridiques précisent les conditions de réutilisation :

- **Les Conditions générales de l'APIE**<sup>46</sup> (octobre 2010). Ces Conditions générales sont une reformulation pédagogique du cadre standard de la loi.
- **Les licences-types de l'APIE**<sup>47</sup>, qui sont des cadres à compléter pour des données la réutilisation est soumise à des conditions particulières (par exemple de mise à jour) et/ou au paiement d'une redevance.
- **La licence OdbL**<sup>48</sup> (dite aussi ODC-ODbL, 2010). Licence internationale établie par l'organisation Opendatacommons. A l'heure où nous écrivons ces lignes, la ville de Paris travaille à une traduction française dont un premier jet peut être consulté et commenté.
- **La licence ODC-by**<sup>49</sup> (2010). Idem.
- **La licence PDDL 1.0**<sup>50</sup> (2010). Idem.
- **Les licences Creative Commons**<sup>51</sup> (depuis 2004). Il s'agit en fait d'un groupe de licences et qui définissent différentes manières pour un auteur de faciliter la circulation et la réutilisation de son œuvre, dans le cadre de la législation sur le droit d'auteur. Son intérêt dans le domaine des contenus littéraires et artistiques n'est plus à démontrer mais ces licences semblent moins bien adaptées au domaine des bases de données. Nous les évoquons à titre d'information et de comparaison.
- **La licence IP du ministère de la Justice**<sup>52</sup> (mai 2010). Une des premières licences type en France, inspirée des Creative commons.

Ces différentes licences partagent certains points communs et se différencient sur d'autres :

### Des fonctionnalités communes

- **L'obligation de la mention d'auteur (BY)** semble être une des fonctionnalités les mieux partagées, puisqu'un seul cadre juridique dispense de la mention d'auteur, la licence PDDL 1.0 – qui par essence dispense de toute obligation.
- La plupart des cadres juridiques **permettent un usage commercial tel quel** : autrement dit, ils rendent possible à un réutilisateur de revendre des données publiques sans les transformer, ni les enrichir. La licence IP du ministère de la Justice fait seule exception et interdit une réutilisation telles quelles de ses données à des fins commerciales : pour tirer un bénéfice commercial de la réutilisation, il faut les enrichir d'une manière ou d'une autre.

### Des fonctionnalités clivantes

Certaines fonctionnalités sont adoptées par certains cadres juridiques et non par les autres. Ces fonctionnalités établissent une sorte de clivage sur quelques grands sujet d'usage des données.

La loi française de 1978 exige, sauf consentement explicite, que les réutilisateurs indiquent la source des données et leur date de mise à jour, et respectent l'intégrité des données. Certaines licences facilitent la levée de telle ou telle de ces obligations :

- **Permettre l'altération des données** est une fonctionnalité explicite des licences de l'Opendatacommons (ODC). Cette fonctionnalité est bien adaptée aux données qui proviennent de sources multiples et dont la mise en valeur exige des traitements. Elle simplifie la réutilisation des données en ce sens qu'un auteur modifiant une donnée n'a pas à en

<sup>45</sup> <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068643&dateTexte=20101103>

<sup>46</sup> <https://www.apiefrance.fr/sections/actualites/des-conditions-generales-pour-la-reutilisation-des-informations-publiques/view>

<sup>47</sup> [https://www.apiefrance.fr/sections/acces\\_thematique/reutilisation-des-informations-publiques/licence-type/view](https://www.apiefrance.fr/sections/acces_thematique/reutilisation-des-informations-publiques/licence-type/view)

<sup>48</sup> <http://www.opendatacommons.org/licenses/odbl/>

<sup>49</sup> <http://www.opendatacommons.org/licenses/by/>

<sup>50</sup> <http://www.opendatacommons.org/licenses/pddl/>

<sup>51</sup> <http://creativecommons.org/about/licenses>

<sup>52</sup> [http://www.rip.justice.fr/information\\_publique\\_liberement\\_reutilisable](http://www.rip.justice.fr/information_publique_liberement_reutilisable)

demander l'autorisation au cas par cas. C'est le cas par exemple des données d'OpenStreetMap : la carte étant enrichie et modifiée en permanence par des dizaines de milliers de contributeurs, il fallait un mécanisme simple et rapide pour laisser chaque contributeur enrichir la carte à loisir.

La loi française de 1978 exige, sauf consentement explicite, que les réutilisateurs indiquent la source des données et leur date de mise à jour, et respectent l'intégrité des données. Certaines licences facilitent la levée de telle ou telle de ces obligations.

- **L'obligation de mention de la date de dernière mise à jour des données** n'est pas prévue par les licences de l'Open Knowledge Foundation. Cela dit, ces licences ne sauraient remplacer la loi et l'on peut penser qu'un réutilisateur de données publiques sous licence ODbL, par exemple, devra tout de même se conformer à cette obligation. Les licences de l'OKFN gagneraient peut-être en efficacité, en France, à expliciter clairement ce point.

D'autres fonctions différencient également certaines licences :

- **L'interdiction de l'usage commercial** des données est une fonctionnalité d'une des licences Creative Commons (CC-NC), la seule à l'adopter. Les données concernées par cette obligation ne peuvent pas être réutilisées dans un cadre commercial. Cependant, la réutilisation des données publiques remplit très largement un but économique : la loi ne permet donc pas, sauf exception qui doit être motivée par l'intérêt général, d'interdire de manière générale les réutilisations commerciales.
- La licence IP propose une fonctionnalité originale en **permettant un usage commercial conditionné par l'obligation d'enrichissement des données**.
- **L'obligation de fournir les données sans entrave technique**, comme par exemple des systèmes techniques de gestion des droits de propriété intellectuelle (DRM), est une fonctionnalité originale de la licence ODbL.

Enfin, les licences-types de l'APIE sont, dans ce paysage, les seules à recouvrir le **paiement d'une redevance** sur les données publiques. La question de la tarification des données est abordée au paragraphe suivant.

### **2.2.3. Faut-il tarifer la réutilisation des données publiques ?**

Les données publiques doivent-elles être mises à disposition des réutilisateurs (qui peuvent être des entreprises installées, mais aussi des "jeunes pousses", des associations, des citoyens, des chercheurs) de manière gratuite ou payante ? Et lorsqu'elles sont payantes, sur quelles bases ?

#### ► **Dans les pays anglo-saxons, le choix de la gratuité**

Au Royaume-Uni et aux Etats-Unis, la question est pour l'essentiel tranchée en faveur de la gratuité, qui ne fait pratiquement plus débat. Plusieurs arguments ont conduit les acteurs publics à ce choix :

- Un argument de principe : la production de ces données est financée par l'impôt. Cet argument comporte deux faiblesses : d'une part, la mise à disposition en dehors de l'administration peut entraîner des coûts ; et d'autre part, les transférer aux entreprises pourrait s'assimiler à une forme de "privatisation des profits", tandis que les coûts demeurent publics.
- Deux arguments économiques : d'une part, le coût de distribution en ligne de la plupart des données publiques est nul, ou presque – et par conséquent son prix "optimal" est très faible, inférieur au coût qu'entraînerait le fait de facturer les réutilisateurs, relancer les mauvais payeurs, etc. D'autre part, les données publiques forment la base d'activités économiques qui contribuent à la croissance et génèrent des rentrées fiscales supérieures aux revenus directs escomptés de la vente des données<sup>53</sup>.

En France, et ailleurs en Europe, la question est plus ouverte. Si les voix en faveur de la gratuité se font entendre, le déséquilibre des comptes publics favorise également la tentation inverse. Comme l'indique ce compte-rendu d'un conseil des Ministres de juin 2010<sup>54</sup> portant sur la revue générale des politiques publiques, "*La mise en œuvre d'une politique volontariste de licences, notamment dans le cadre de leur réutilisation commerciale, peut rapidement augmenter les produits pour l'Etat et ses établissements publics. En effet, les licences déjà octroyées qui concernent, à l'heure actuelle, essentiellement les secteurs traditionnels de l'information géographique, météorologique, économique et financière, rapportent chaque année environ 50 millions d'euros à l'Etat.*"

#### ► **Ce que dit la loi**

La loi prévoit que les détenteurs des données publiques peuvent exiger une redevance dans le but de couvrir les coûts de collecte, de production et de mise à disposition, ainsi que les investissements consentis. Ce principe établit un cadre qui peut sembler contraignant, mais laisse pour l'instant assez nombreuses zones d'ombre :

- Le but ne peut pas être de dégager des bénéfices sur la mise à disposition des données publiques. Mais que revêt la couverture des coûts ? Si des données doivent de toute manière être produites dans le cadre normal de l'activité de service public (financée par l'impôt), il paraît difficilement justifiable d'en faire payer autre chose que les coûts spécifiquement liés à leur mise à disposition à des fins de réutilisation – coûts qui seront souvent assez faibles, du moins si les systèmes informatiques ont été conçus comme il faut.
- La rémunération doit respecter les règles de la concurrence et être "équilibrée, proportionnelle et non discriminatoire". L'exclusivité est interdite, ainsi que la discrimination entre deux acteurs ayant des besoins équivalents. Un réutilisateur public ne peut par exemple pas être favorisé par rapport à une entreprise, du moins pour le même type de réutilisation. Il est cependant possible d'établir des différences, par exemple entre usage commercial et non-commercial, ou encore entre des "gross" réutilisateurs et des "petits".

---

<sup>53</sup> En 2008, une étude préalable avait été commandée par le gouvernement britannique à 3 chercheurs de l'université de Cambridge (Pollock, Newbery, Bentley 2008). L'Etat devait-il, du point de vue de l'efficacité économique, donner ses données ou les faire payer ? Selon leurs conclusions, "la chose la plus intelligente à faire pour l'Etat est de faire payer le coût de distribution, ce qui tend vers zéro pour une distribution en ligne et donc revient à les mettre à disposition gratuitement". Cette analyse des coûts et bénéfices du modèle commercialisé a donc conclu en faveur d'une mise à disposition gracieuse.

<sup>54</sup> <https://www.apiefrance.com/sections/actualites/conseil-des-ministres-du-30-juin-2010-revue-generale-des-politiques-publiques/view>

En revanche, il est clair que ce sont les coûts qui doivent guider la tarification, et non un objectif de maximisation du revenu direct susceptible d'être tiré de la mise à disposition des données.

## ► Comment définir sa stratégie de tarification ou de gratuité ?

Plusieurs facteurs doivent être pris en compte<sup>55</sup> :

- **Combien coûte réellement la mise à disposition des données – et combien coûte le fait de les facturer ?** Pour un système informatique correctement conçu, qui sépare efficacement "données" et "traitements" (un principe appliqué depuis plusieurs décennies), la réponse pour des fichiers relativement simples et pérennes est clairement : rien, ou presque – en tout cas bien moins que le coût de facturation et de recouvrement. En revanche, s'agissant de bases de données complexes ou encore de données en temps réel, il ne s'agit plus d'exporter des fichiers mis en téléchargement, mais d'ouvrir des "portes d'entrée" sur les infrastructures informatiques de l'acteur public, qu'il aura souvent fallu adapter à cet effet. L'investissement peut s'avérer significatif, et si la demande est forte, le coût d'exploitation ne sera pas non plus négligeable<sup>56</sup>.
- **Quel est le consentement à payer, de la part de qui ?** Les grands acteurs du "marché de l'information" seront prêts à payer plus que les petites entreprises, et *a fortiori* que les associations – y compris, d'ailleurs, pour évincer les nouveaux entrants. On payera plus volontiers des données enrichies que des données brutes. Bref, tout dépend de ce que les acteurs publics veulent privilégier, ainsi que de *qui* ils veulent privilégier.
- **Quels sont les objectifs de long terme de l'acteur public ?** Celui-ci peut en effet poursuivre plusieurs objectifs différents, et pas toujours compatibles :
  - Développer des revenus extrafiscaux, en facturant le plus possible. Cet objectif invite à facturer au plus près du "consentement à payer", en flirtant avec les limites de la loi.
  - Favoriser le développement de l'activité économique, de l'innovation sociale et des connaissances que peuvent induire les données publiques. C'est ce que visent les Anglo-saxons, c'est également la raison pour laquelle un puissant mouvement (dans lequel la France se montre active) se développe en faveur de la gratuité des données scientifiques. Cet objectif invite plutôt à la gratuité, sauf dans le cas où le coût de mise à disposition est très élevé.
  - Atteindre des objectifs politiques plus larges, en matière d'éducation, de recherche, de culture, de développement durable, de participation démocratique, etc. Cette orientation tend également vers la gratuité, du moins dans les domaines considérés comme prioritaires.
- **Enfin, comment éviter que la grille tarifaire elle-même ne dissuade la demande ?** C'est le cas en particulier lorsqu'un dispositif trop "intelligent" crée une complexité incompréhensible pour la plupart des acteurs. Une administration nationale a par exemple récemment diffusé une grille tarifaire dans laquelle elle distingue une demi-douzaine de cas (et en leur sein, une multitude de sous-cas), dont certains ne peuvent que créer une incertitude totale pour les réutilisateurs. S'agissant des entreprises, elle distingue par exemple entre un usage "interne" et un usage "externe" (commercial) ; quant à l'usage externe, le tarif est proportionnel au chiffre d'affaires lié à la réutilisation des données, alors que cette réutilisation se fait très souvent en combinant plusieurs données de plusieurs sources, rendant impossible la mesure de la contribution d'une donnée précise...

Sans faire de la gratuité un principe, on ne saurait trop inviter les acteurs locaux à ne tarifer leurs données qu'en cas de réelle nécessité, après avoir exploré toutes les autres pistes. Dans la plupart des cas (au moins dans les premiers temps), les inconvénients de la tarification dépassent ses avantages.

<sup>55</sup> Pour aller plus loin, on pourra se référer à l'étude économique réalisée en 2010 par le Bureau d'Economie Théorique et Appliquée (BETA) de l'université de Strasbourg pour le compte de l'APIE, mais non sans en exercer une lecture critique. Le BETA part en effet de l'idée que la mise à disposition de données sous une forme "brute" (même documentées) en limite très fortement la réutilisation, alors que la plupart des réutilisateurs, même petits, semblent réclamer des données aussi brutes que possible.

[https://www.apiefrance.fr/sections/acces\\_thematique/reutilisation-des-informations-publiques/etude-economique/view](https://www.apiefrance.fr/sections/acces_thematique/reutilisation-des-informations-publiques/etude-economique/view)

<sup>56</sup> Sur son service Google Maps, Google a tranché d'une manière intéressante : les petits utilisateurs ne sont pas facturés, tandis que ceux qui consomment une part mesurable des ressources informatiques le sont.

## 2.3. COMMENT RENDRE LES DONNEES ACCESSIBLES ?

La diffusion des données est une étape clé du processus de partage des données publiques. Face à ces problématiques qui comportent parfois des dimensions assez techniques, nous présentons ici les grands traits de ce sujet, pour y voir plus clair et pour commencer à avancer sans attendre que tout soit résolu.

### Le W3C s'intéresse à la publication de données publiques

Le W3C, l'organisation internationale qui édicte les standard du web – la norme HTML, c'est lui ! – a publié en septembre 2009 un "Working Draft" (document de travail) sur la publication des données publiques : [Publishing Open Government Data](#)<sup>57</sup>. Si le processus de standardisation de ce document en est encore à ses débuts, ce document pourra néanmoins intéresser les lecteurs avancés qui voudraient anticiper à sa rédaction, voire y participer.

### 2.3.1. Formats et méthodes d'accès aux données

(note : les non-techniciens pourront se contenter de lire les "grands principes" avant de passer à la partie suivante)

#### ► Les grands principes

Pour être facilement réutilisables, vos données doivent partager quelques règles simples.

#### Des données documentées

Les données doivent d'abord être documentées, en ce qui concerne leur contenu, leur format, leurs moyens et leurs conditions d'accès.

Il s'agit d'une étape importante car elle conditionne la réutilisation pertinente de vos données. Une donnée n'est jamais trop décrite, expliquée.

Prenons par un exemple un champ intitulé "nom" : s'agit-il d'un nom de personne, d'un nom de produit, du nom d'un auteur, d'un fabriquant, d'un usager, ou d'autre chose encore ? Prenons un exemple apparemment moins problématique, un champ "date". S'agit-il d'une date de création, de changement, d'une date approximative, d'une date précise, d'une date de modification de la donnée ? Et comment cette date s'écrit-elle : 2010-12-02 ? 02/12/2010 ? 12-dec-2010 ? La documentation des données facilite leur compréhension et donc leur réutilisation.

La documentation concerne tous les niveaux de la diffusion des données :

- La sémantique des données : que représente cette donnée, que signifie chaque champ d'un fichier ?
- La syntaxe des données : comment s'écrit chaque information ?
- Le format des fichiers contenant les données
- La méthode d'accès aux données

Pour chaque donnée on tâchera de préciser les éléments suivants :

- Une description libre en bon français tâchant de décrire au mieux chaque donnée
- L'éventuel format ou norme formelle adoptée (par exemple pour une date)
- Eventuellement la taille maximum ou fixe de la donnée
- Le type de donnée : alphanumérique, numérique, binaire, etc.

On précisera si le schéma global des données est lui-même normalisé ou repris à l'identique d'une autre base.

Enfin, il est important de documenter le jeu de données en lui-même : d'où proviennent ces données ? Sont-elles brutes ou issues du traitement de données "primaires" – et dans ce cas,

<sup>57</sup> <http://www.w3.org/TR/gov-data/>

lesquelles ? Quelle est leur dernière date de mise à jour ? Leur fréquence de mise à jour ? Qui en est l'auteur ? Y a-t-il des conditions de réutilisation, des licences ? Etc.

### **Données brutes ou retraitées ?**

Les réutilisateurs revendentiquent la plupart du temps des "**données brutes**", pour une meilleure transparence. Ils peuvent aussi apprécier des données "retravaillées", "enrichies", à partir du moment où les traitements sont correctement documentés.

Dans la plupart des cas, on aura intérêt à livrer les données brutes par défaut, quitte à proposer également des données plus élaborées.

### **Privilégier les standards ouverts**

Enfin, au-delà de la documentation et de la diffusion de données brutes, le producteur de données regardera les **standards ouverts** relatifs aux données diffusées. Certains domaines ont vu naître de tels standards qui concourent à une meilleure interopérabilité et diffusion des savoirs. Par exemple, le fameux format RSS, servant de format de syndication de contenu web, est devenu un incontournable. Il serait dommage de réinventer la roue...

Un format ouvert est caractérisé par "des spécifications techniques documentées, publiées, non payantes, sans brevet dessus, sans royalties dessus" pour des données dont l'usage est "indépendant d'un logiciel particulier, d'un système d'exploitation ou d'une société"<sup>58</sup>. Un format ouvert garantit donc la maîtrise des données, leur pérennité, leur archivage et leur interopérabilité – et facilite donc la réutilisation.

#### **Standard ouvert : une définition juridique pour aider les acteurs publics**

Le terme de "standard ouvert" est défini par la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique, comme suit : "tout protocole de communication, d'interconnexion ou d'échange et tout format de données interopérables et dont les spécifications techniques sont publiques et sans restriction d'accès ni de mise en œuvre" (article 4).

Dans ses relations avec ses fournisseurs technologiques, un acteur public peut avoir intérêt à exiger la livraison d'outils qui produisent des données selon un standard ouvert, au sens de cette loi.

### ► **La sémantique des données**

La documentation de la sémantique ("que signifie tel champ d'un fichier ?") est parfois négligée car elle peut paraître évidente au producteur des données. Même une simple date peut poser un problème de sémantique : s'agit-il du calendrier grégorien ou d'un calendrier musulman ? L'encyclopédie Wikipedia recense plus d'une trentaine de calendriers<sup>59</sup>...

#### **Normaliser la sémantique des données sur le web : le modèle RDF**

Le format RDF<sup>60</sup> permet de documenter, de décrire de façon standard des jeux de données sur le web en utilisant des propriétés publiques et non ambiguës, regroupées en "vocabulaires". Par exemple, le "vocabulaire RDF" appelé FOAF décrit la propriété "name" comme le nom propre d'une personne. Pour être sûr qu'il n'y ait pas d'ambiguïté, cette propriété "name" est elle-même documentée à une adresse web unique (URI) contrôlée par l'auteur du vocabulaire FOAF : <http://xmlns.com/foaf/0.1/name>.

Le format RDF, en plus d'être un format ouvert, améliore l'interopérabilité des données. Il est à la base de ce qu'on appelle le "web sémantique" ou "web des données".

<sup>58</sup> <http://formats-ouverts.org/blog/2004/07/01/12-un-article-de-loi-definit-ce-que-sont-les-formats-ouverts>

<sup>59</sup> <http://fr.wikipedia.org/wiki/Calendrier>

<sup>60</sup> <http://www.w3.org/RDF/>

## ► La syntaxe des données

La syntaxe des données doit aussi faire l'objet d'une attention particulière. Beaucoup d'éléments codifiés peuvent l'objet de syntaxes très différentes qui, si on ne les connaît pas, peuvent altérer la compréhension des informations. On l'a vu plus haut pour l'exemple des dates mais cela vaut aussi pour de très nombreuses autres données : ISBN, ISSN, numérotations en tous genres (numéros de département, codes postaux, numéros INSEE...), etc. Il est donc utile de documenter clairement la syntaxe de chaque champ.

Là encore, il n'est pas rare de voir des standards ouverts couvrant la syntaxe de certaines données.

## ► Les méthodes d'accès aux données

Il existe deux façons de donner accès à des données sur le web : le téléchargement d'un fichier, ou l'accès direct à la source de la donnée.

### Les fichiers téléchargeables

L'accès direct à un fichier contenant les données est souvent la méthode la plus simple car les outils informatiques proposent généralement une fonction d'export des données sous forme de fichier. Elle est cependant moins bien adaptée lorsque les données changent souvent et qu'il y en a beaucoup : cela oblige les réutilisateurs à recharger fréquemment l'ensemble du fichier et peut être consommateur en ressources système et réseau.

### Les interfaces de programmation (API)

Cette limite explique pourquoi se développe fréquemment un mode d'accès direct aux données uniquement nécessaires à travers une "interface de programmation" (API en anglais). Une API va permettre à une application "demandeuse" (celle du réutilisateur) de demander à l'application "fournisseur" (celle du détenteur des données "sources") les informations dont elle a besoin. Les grands acteurs du numérique en général et du web particulier, comme Google, Amazon, Facebook, etc., ont établi leur succès grâce à leurs APIs : leurs données sont en quelque sorte "réutilisables", sans que leurs bases de données ne quitte jamais leurs ordinateurs.

Il existe deux grands types d'APIs dans le monde du web : les APIs REST et SOAP.

- SOAP est riche en fonctionnalités mais complexe à mettre en œuvre, en particulier pour les réutilisateurs.
- A l'inverse, les APIs REST sont plus légères et plus simples à mettre en œuvre. Ses dernières sont d'ailleurs plébiscitées par les producteurs de données comme par les réutilisateurs. SOAP connaît quelque succès dans les réseaux internes des grandes entreprises mais il est presque inexistant sur le web : la plupart des grands acteurs du web ou des grands producteurs de données publiques ne l'utilisent pas.

Pour être efficace, l'API doit être clairement documentée. Par ailleurs, selon vos jeux de données, il est possible de choisir une API existante plutôt que de réinventer la roue. Si vous utilisez des données décrites en RDF, il peut être intéressant de mettre en œuvre l'API standard SPARQL<sup>61</sup>. SPARQL est à la fois un protocole et un langage de requête pour les données en RDF. Il permet d'interroger toutes les données d'une base de données accessible sur le web à l'aide d'une seule technique : il n'est donc pas nécessaire de développer une API pour chaque type de données. En outre, il existe déjà de nombreux outils de diffusion et d'interrogation qui reposent sur SPARQL, évitant ainsi un travail fastidieux.

### Une API standard pour le suivi d'anomalies impliquant les services publics du territoire ?

<http://open311.org>

Accès au services d'enlèvement des objets encombrants, signalement d'éclairages défectueux ou de graffitis, lieux et horaires des centres de tris : autant d'activités que les territoires traitent aujourd'hui, souvent, à l'aide de centres d'appels dédié : "Allo mairie", "Allo propreté", etc. Open311, un nom inspiré par le numéro de téléphone standard pour ce genre de services aux États-Unis – le 311 –, est une API ouverte qui permet de standardiser les données et les

<sup>61</sup> <http://fr.wikipedia.org/wiki/SPARQL>

échanges de données dans le suivi d'anomalies impliquant les services publics d'un territoire.

Cette API est d'ores et déjà utilisée par quelques grandes villes américaines comme New York ou San Francisco. Elle permet, non seulement d'accéder à l'information, mais également d'en envoyer : déclarer un problème, par exemple. "En rendant publique cette information, cela produit de la transparence et une meilleure prise en compte de la part de ceux qui ont la charge de ces problèmes. La transparence garantit aussi que la voix de tous est écoutée et encourage à plus de participation<sup>62</sup>.", indique le site d'Open311.

Cette API rencontre un certain succès avec plus d'une vingtaine d'applications qui l'exploitent pour des usages parfois très originaux : la cartographie du bruit dans la ville, un tableau de bord hyperlocal (l'activité publique autour de mon pâté de maison), des jeux participatifs, etc.

### ► Dans quel format de fichier publier ses données ?

"N'importe quel format", demandent de nombreux réutilisateurs, "du moment qu'il est ouvert et documenté". Les standards ouverts sont importants, ils sont toujours préférables, mais un fichier bien documenté qui ne respecte aucun standard est toujours préférable à pas de fichier du tout. L'absence de standard ne constitue pas une bonne raison d'attendre...

#### Les formats de fichiers

Nous l'avons dit, les formats ouverts garantissent une meilleure réutilisabilité des données. Certains formats de fichiers propriétaires peuvent cependant être acceptables lorsqu'ils sont extrêmement répandus et qu'on trouve facilement des outils à bas coût pour les manipuler. Par exemple, le format .xls (pour les documents issus du tableur Excel de Microsoft), s'il n'est pas ouvert, peut éventuellement convenir et sera toujours "mieux que rien". A coût de production égal, on lui préfèrera cependant le format Open Document<sup>63</sup> qui est, lui, réellement ouvert.

#### Les formats d'API

Dans le monde des API web, les formats ouverts XML et JSON<sup>64</sup> règnent en maîtres.

- On ne présente plus XML et ses qualités dans la gestion des documents structurés, mais beaucoup d'experts considèrent qu'il est plus adapté au traitement de documents que de données.
- Le format JSON, plus récent, est né de ce constat et vise tout particulièrement le traitement des données. Ses promoteurs assurent qu'il présente l'intérêt d'être plus simple et plus compact que XML et de mieux interagir avec les langages de programmation. Depuis 2010, par exemple, le site de microblogging Twitter.com, succès planétaire avec plus de 100 millions de comptes, ne délivre plus de données en XML au profit de JSON. JSON a aussi ses détracteurs et les défauts d'un nouvel entrant face à un XML bien implanté et qui continue de progresser<sup>65</sup>.

### 2.3.2. Informer sur les données disponibles : répertoires et portails

Il ne suffit pas de mettre ses données en ligne : encore faut-il le faire savoir, et permettre à tous ceux qui s'y intéressent de savoir quelles données sont disponibles, sous quelle forme et dans quelles conditions. Il existe plusieurs moyens et plusieurs "espaces" pour informer sur les données disponibles, tous n'étant pas nécessairement créés et gérés par les détenteurs des données.

### ► Une obligation : informer sur les informations publiques que l'on détient

En France, la loi oblige tous les acteurs publics à **informer le public des informations publiques qu'ils produisent ou détiennent**, sous la forme d'un *répertoire des informations publiques*. S'ils disposent d'un site internet, ce répertoire doit alors exister sous une forme électronique.

<sup>62</sup> <http://open311.org/learn/>

<sup>63</sup> <http://fr.wikipedia.org/wiki/OpenDocument>

<sup>64</sup> <http://www.json.org/>

<sup>65</sup> E4X, Ecmascript pour XML, est une extension d'Ecmascript qui rend plus simple la lecture du XML. Elle d'ores et déjà intégrée dans Firefox et Google Chrome : [http://fr.wikipedia.org/wiki/ECMAScript\\_pour\\_XML](http://fr.wikipedia.org/wiki/ECMAScript_pour_XML)

Ce répertoire doit mentionner pour chaque document ou jeu de données, au moins : le titre, l'objet, la date de création, les conditions de réutilisation, la fréquence et la nature des mises à jour.

Ce répertoire concerne *toutes* les informations de service public détenues ou mises à jour par les acteurs publics, qu'elles soient ou non mises à disposition sous la forme de données informatiques accessibles en ligne. Le répertoire facilite à la fois le droit d'accès, et celui de réutilisation.

Mais, bien évidemment, une action délibérée de mise en ligne des données publiques, telle que nous la décrivons dans ce guide, ne sera pas complète sans un répertoire complet, accessible et bien organisé des données.

Dans tous les cas, il est dans l'intérêt de l'acteur public, qu'il s'engage ou non dans une politique volontariste de réutilisation de ses données, de fournir des informations supplémentaires tels que la personne à contacter pour demander l'accès aux données, le mode d'accès aux données, le format technique des données, etc. L'APIE mentionne sur son site plusieurs répertoires à titre d'exemple, et propose un guide technique très complet<sup>66</sup>.

## ► Premiers pas : un répertoire simple

Au-delà de l'obligation légale, un bon **répertoire des informations publiques** peut offrir **une bonne base de publication et de référencement des données** :

- Il liste en principe toutes les données détenues (qu'elles soient en ligne ou non) ;
- Il donne une vue d'ensemble qui permet aux réutilisateurs de comprendre ce qui existe et ce qui est plus ou moins aisément disponible ;
- Il décrit précisément chaque jeu de données ;
- Il peut, soit constituer la plate-forme sur laquelle certains jeux de données seront physiquement disponibles, soit "pointer" sur les sources des données qu'il référence.

Certains acteurs peuvent faire le choix de distinguer le répertoire exhaustif des informations de service public (exigé par la loi de 1978) et le répertoire plus restreint des "données publiques" qu'ils mettent en ligne. Cette option a du sens du point de vue des réutilisateurs. Encore faut-il ne pas confondre deux objectifs :

- Recenser les informations mises en ligne à des fins de réutilisation ne dispense pas de l'obligation légale de répertorier *toutes* les informations de service public détenues ;
- Un répertoire des données publiques mises en ligne doit être pensé comme un service destiné aux réutilisateurs, pas comme une simple liste.

## ► Du répertoire au "portail"

La plupart des acteurs publics qui ont ouvert leurs données en France ont ainsi choisi de créer une **plateforme ou un portail de publication spécifique**. Ces plateformes sont conçues dans le but de cataloguer et diffuser des données réutilisables. Elles vont à l'essentiel sans perdre de temps à lister ce qui n'est pas disponible. Par ailleurs, ces plateformes intègrent des services qui facilitent la réutilisation. Souvent il s'agit :

- De "métadonnées" précises destinées à informer les réutilisateurs : description des ressources, conditions de réutilisation, contacts...
- De la documentation technique précise des méthodes d'accès (fichiers ou APIs)
- De forums de discussion pour aider les réutilisateurs dans leurs démarches
- D'espaces de promotion des usages et des applications réalisées grâce aux données

Comme souvent pour ce genre d'outil, il existe des solutions de plateformes/portail très différentes, qui répondent aux différents besoins des acteurs publics.

Certains acteurs pourront choisir un **développement spécifique (ou sous-traité)**, comme c'est le cas de certains grands portails nationaux aux États-Unis, en Australie, et bientôt en France. Un développement spécifique peut être utile dans plusieurs cas : par exemple quand l'acteur public doit intégrer des fonctionnalités spécifiques correspondant à son environnement ou plus simplement quand il ne trouve pas d'autres solutions adaptées à ses besoins et ses contraintes. Il présente en revanche l'inconvénient de coûter cher, de prendre du temps et

---

<sup>66</sup> <http://bit.ly/a5rRk0>

d'aboutir à un dispositif difficile à faire évoluer. Cette voie n'est pas recommandée aux acteurs territoriaux.

Depuis 2010, plusieurs initiatives de **plateformes "open source"** permettent de monter un portail à peu de frais :

- Après l'avoir développée à l'origine pour Rennes Métropole, la société In-Cité a publié récemment sa solution de plateforme libre sous licence "libre" (GPL) : <http://opendata.incite.net/>. Ce projet, basé sur un outil très répandu de gestion de contenus web (Typo3), fournit une solution complète pour acteur public, qui inclut le catalogage des données, leur diffusion, ainsi que des outils de dialogue avec les réutilisateurs.
- Le logiciel libre CKAN<sup>67</sup>, créée sous l'égide de l'*Open Knowledge Foundation* britannique, se focalise quant à lui sur la partie catalogage et diffusion des jeux de données<sup>68</sup>. Il connaît déjà un certain succès puisqu'il motorise le portail officiel du gouvernement [data.gov.uk](http://data.gov.uk) ainsi que plusieurs autres sites anglais, allemands, canadiens et français.

Enfin, il est aussi possible de faire appel à un acteur tiers qui créera et hébergera le portail sur sa propre infrastructure informatique. Là encore on distingue plusieurs cas de figures possibles :

- Les **plateformes privées** en ligne sont encore rares. En France nous avons la chance d'avoir un projet riche et abouti : Data Publica ([www.data-publica.com](http://www.data-publica.com))
- Si l'offre n'est pas encore disponible, ces tous prochains mois vont voir apparaître les **plateformes publiques multi-acteurs** : elles sont opérées par des acteurs publics pour le compte d'autres acteurs publics. Le projet le plus significatif est le futur portail "Etatlab" de l'APIE qui doit voir le jour en 2012. Certaines collectivités publiques, comme le département de la Gironde, sont en train de réfléchir à un tel modèle.

Enfin, si cette offre n'est pas non plus disponible actuellement, nous savons qu'il existe chez plusieurs grands intégrateurs informatiques, des réflexions sur des **plateformes en "marque blanche"**. Il s'agit d'outils génériques, conçus pour pouvoir s'adapter facilement aux besoins d'une collectivité publique particulière.

## ► **Les portails indépendants et "portails de portails"**

Enfin, on assiste également à l'émergence d'initiatives publiques, privées et associatives qui visent à répertorier des sources de données provenant de plusieurs acteurs :

- Le portail français "Etatlab" recensera des données provenant de l'ensemble des acteurs publics nationaux, mais n'exclut apparemment pas d'intégrer également des données provenant, par exemple, de collectivités locales.
- Le site français Data Publica poursuit le but de publier ou référencer n'importe quel type de donnée ouverte, y compris celles venant d'acteurs privés, à travers de nombreux mécanismes de publication et d'échange.
- Le projet CKAN<sup>69</sup> – pour *Comprehensive Knowledge Archive Network* – de l'Open Knowledge Fondation s'est donné pour mission de favoriser la diffusion et la réutilisation automatique de jeux de données à travers le monde. Il s'agit d'un registre "ouvert" qui référence et mutualise des données ou toute autre source de connaissances. Une application facilite la recherche, le partage et la réutilisation de données.

De son côté, le collectif [Regards Citoyens](#) a ouvert en novembre 2010 un portail de recensement des données publiques librement accessibles en France : [nosdonnees.fr](http://nosdonnees.fr). Le site propose pour l'instant un registre libre de jeux de données ainsi qu'une "plateforme collaborative de chasse aux trésors", qui vise à identifier les données publiques disponibles même lorsque celles-ci sont difficiles à identifier. Ce portail s'appuie sur le logiciel CKAN, en l'adaptant au contexte français (langue et jeux de données).

<sup>67</sup> <http://okfn.org/projects/ckan/>

<sup>68</sup> <http://knowledgeforge.net/ckan/trac/#Features>

<sup>69</sup> [www.ckan.net/](http://www.ckan.net/)

### Datalift : un projet de recherche de plateforme d'interconnexion de données

[Datalift](#)<sup>70</sup> est un projet de recherche expérimentale financé par l'agence nationale de la recherche. Son but est de développer une plateforme pour publier et interconnecter des jeux de données sur le web de données. Datalift tout à la fois publie des jeux de données provenant d'un réseau de partenaires et propose un ensemble d'outils facilitant le processus de publication de jeux de données. L'interconnexion des jeux de données va permettre de rapprocher des données qui ne l'étaient pas au départ, pour des usages toujours plus riches.

<sup>70</sup> <http://datalift.org/>

## Les portails britanniques de recensement de données publiques

Le portail gouvernemental [data.gov.uk](http://data.gov.uk) dresse la liste des données publiques ouvertes. Actuellement il héberge ou référence environ 3200 bases de données et une cinquantaine d'applications dérivées. Il référence plus d'informations que le portail américain [data.gov](http://data.gov). Le portail britannique se différencie aussi de son homologue américain en ayant fait le choix de formats standardisés favorisant le développement du web sémantique<sup>71</sup>. Les données peuvent être recherchées par thèmes, par mots-clés ("tags"), par sources, par régions géographiques, etc. La gestion du portail est partagée entre le cabinet du Premier ministre et les archives nationales.

The screenshot shows the homepage of data.gov.uk. At the top, there's a navigation bar with links for Data, Apps, Ideas, Forum, Wiki, Blogs, Linked Data, Resources, and About. A "Log in or sign up" button is also present. Below the navigation is a search bar with the placeholder "What are you looking for?". The main content area features a large "Data" section with a search bar. To the right, there's an "EDITOR'S PICK" section for "COINS – HM Treasury's database of public spending". Below that is a "Popular tags" section with various categories like health, care, population, and transparency. Further down, there are sections for "Featured data" and "Random data set", each displaying a preview of a dataset with its title, description, and rating.

Le portail [Openlylocal](http://openlylocal.com)<sup>72</sup>, porté par l'association Countculture, entend donner une meilleure visibilité et faciliter l'accès aux données publiques locales. Chris Taggart et son équipe mènent un travail de recensement des données existantes au sein des *councils* britanniques (plus de 140 d'entre eux étaient "couverts" au début 2011), auquel s'ajoute un travail de mise en forme des données, pour que celles-ci puissent être facilement réutilisées et croisées.

The screenshot shows the homepage of openlylocal.com. At the top, there's a header with the logo "openlylocal making local government more transparent" and navigation links for ALL COUNCILS, HYPERLOCAL SITES DIRECTORY, OPEN DATA SCOREBOARD, and COUNCIL SPENDING. A search bar is also present. The main content area includes a news section titled "Openly Local news" with a link to "Not the way to build a Big Society: parts NESTA". Below that is a section titled "We currently have information on:" listing various types of data available. There's also a "Most recently updated councils" section and a "Latest councillors" section. On the right side, there's a "Forthcoming meetings" section listing various council meetings. A sidebar on the right contains a news item from "Openly Local news" about NESTA and a section titled "Opening up council accounts... and open procurement...".

<sup>71</sup> Cf. Claire Goulard "Les bons ingrédients pour une ouverture des données réussie", Ovni - <http://owni.fr/2010/05/31/opendata-12-data-gov-ou-data-gov-uk/>

<sup>72</sup> <http://openlylocal.com/>

## ANNEXES

---

### A1. "L'OUVERTURE DES DONNEES ET APRES..."

---

Article de Daniel Kaplan, Fing (octobre 2010)

[www.internetactu.net/2010/11/09/louverture-des-donnees-publiques-et-apres](http://www.internetactu.net/2010/11/09/louverture-des-donnees-publiques-et-apres)

**Tous ceux qui, comme nous, s'engagent en faveur de l'ouverture et de la réutilisation des données publiques, en espèrent des résultats féconds en termes de qualité de vie, de cohésion sociale, d'innovation et de croissance. Mais les choses pourraient se passer tout autrement. Nous devons commencer à penser aux conséquences de l'ouverture des données, pour nous assurer qu'elles soient majoritairement positives.**

Ceux qui militent en faveur de l'ouverture des données publiques (ou non, d'ailleurs) et de leur réutilisation par les citoyens, les chercheurs et les entrepreneurs, espèrent qu'il en sortira quelque chose de bon. Ils ont sans doute une idée différence de ce qu'est ce "quelque chose", et de ce que "bon" veut dire : je peux considérer qu'il est bon de créer de la valeur marchande à partir de données publiques gratuites, alors que d'autres ne le penseront pas. Je peux détester les "crime maps" qui se multiplient dans les pays anglo-saxons, parce qu'elles stigmatisent sans rien résoudre, mais d'autres espèrent qu'elles sauveront des vies. Jusqu'ici, tout va bien. C'est même une bonne raison de soutenir l'ouverture des données : parce qu'elle crée le terreau commun sur lequel différents acteurs, avec des motivations différentes, créeront des choses différentes – et Dieu, ou Darwin, reconnaîtra les siens au bout du compte.

#### Explorer les conséquences

Mais les choses ne s'avéreront pas forcément aussi simples. L'ouverture des données publiques est un processus vaste, complexe, sans fin, qui implique des milliers d'acteurs. Elle redistribue l'information, le pouvoir, les responsabilités, selon des lignes difficiles à anticiper. Ses conséquences se feront probablement sentir dans de très nombreux domaines : les métiers des opérateurs de services, les logiciels de gestion, la démocratie participative, les fiches de postes des fonctionnaires, le budget des organismes publics spécialisés dans la production de données (comme l'IGN ou l'Insee), l'économie et le contenu des médias, etc.

Et, selon la manière dont nous ouvrirons l'accès aux données, et selon ce que nous ferons après avoir ouvert cet accès, certaines de ces conséquences pourraient être moins positives que nous ne le souhaiterions.

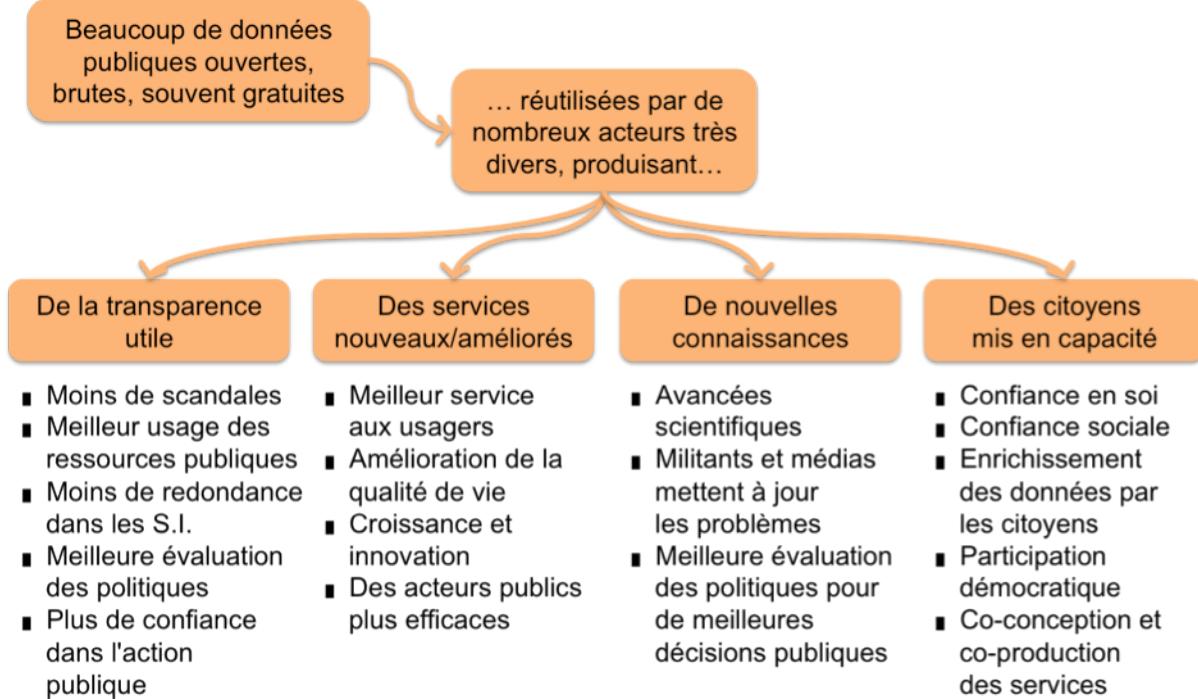
Commençons donc à penser aux conséquences. Imaginons que nous avons gagné : une part très significative des "données de service public" sont désormais accessibles et réutilisables, brutes, en un format lisible par des machines, à un coût faible voire (le plus souvent) nul. Avançons de quelques années : en quoi cela a-t-il changé la vie quotidienne des citoyens et des entreprises ? Et notre potentiel d'innovation et de croissance ? Et la capacité des acteurs publics à atteindre des objectifs collectifs (environnementaux ou sociaux, par exemple) ? Et la démocratie, l'esprit public ?

#### Un avenir radieux

Le récit dominant de l'ouverture des données publiques décrit un avenir radieux. Des groupes de citoyens mettent à jour les abus de pouvoir, et participent de manière active à des débats publics où l'échange d'avis informés permet d'aboutir à de meilleures décisions. Ils contribuent même à enrichir le stock des données publiques, comme ils le font déjà aujourd'hui sur Open Street Map. Des grandes entreprises comme des jeunes pousses créent toutes sortes de nouveaux services et contribuent à la croissance et à la qualité de la vie. Chercheurs et "data-journalistes" passent en revue des masses de données pour enrichir la compréhension, par exemple, des changements climatiques ou des dynamiques urbaines. Les institutions publiques collaborent entre elles pour

éliminer des procédures redondantes ; et avec les entreprises et les associations pour produire de meilleurs services publics. Nourrie par un aller-et-retour constant d'informations, d'évaluations et d'échanges, la démocratie trouve une nouvelle jeunesse.

### **Données publiques ouvertes : le récit dominant d'un avenir radieux**



Nous pouvons bien sûr produire plusieurs exemples et quelques études à l'appui de ce récit. Mais si nous décidons au contraire de nous concentrer sur le côté sombre, il devient assez facile, à *partir du même point de départ*, de raconter une histoire tout à fait différente.

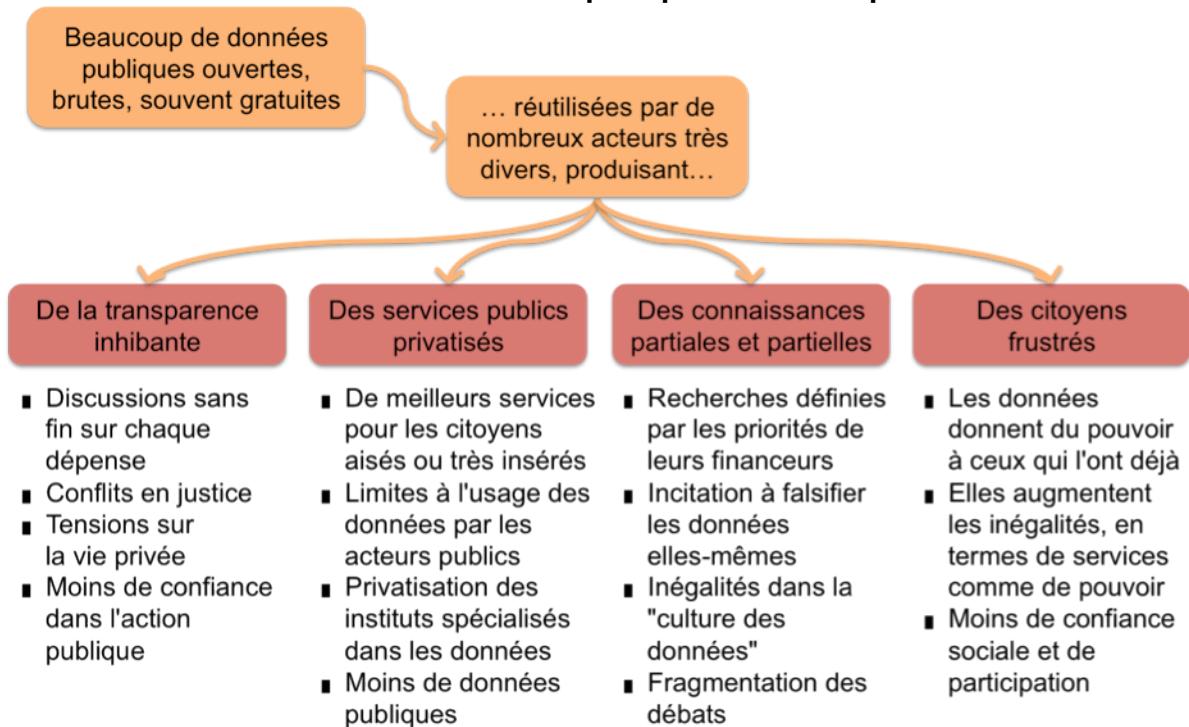
### **Un tour du côté sombre**

Supposons, par exemple, que la plupart des services innovants fondés sur l'usage de données publiques sont proposés par des entreprises, et financés soit par leurs utilisateurs, soit par la publicité. Pour vivre, ces entreprises se focalisent sur les territoires et les populations les plus solvables. Après avoir montré que leurs services sont bien meilleurs que ceux des institutions, elles obtiennent que celles-ci n'aient plus le droit de leur faire concurrence. Le secteur public se retrouve cantonné à servir les populations délaissées par le marché, et les inégalités devant les services fondamentaux s'accroissent.

Supposons ensuite que les instituts publics dont le métier consiste à produire des données, géographiques ou statistiques par exemple, ne parviennent plus à tirer de revenus significatifs de la vente de ces données. Elles tentent, sans y parvenir, de concurrencer les cabinets d'études ou les éditeurs professionnels. L'état des finances publiques ne permettant pas d'augmenter leur dotation budgétaire, elles doivent réduire leurs activités. Beaucoup de données cessent d'être produites, ou bien le sont par des entreprises. L'ouverture des données publiques réduirait ainsi le stock de données publiques...

La transparence peut également produire des conséquences problématiques. On peut imaginer qu'au lieu de s'engager dans des discussions ouvertes et constructives sur des sujets d'intérêt général, les groupes d'intérêt s'en servent pour contester la moindre décision, la moindre ligne de dépense publique. Au lieu d'aboutir à des décisions plus sages et plus largement soutenues, la transparence totale dresse les intérêts particuliers les uns contre les autres, inhibe la prise de décision et réduit la confiance. La nouvelle visibilité des données invite même à les trafiquer à la source, pour faire en sorte que même des analyses indépendantes, fondées sur des moyens informatiques massifs, produisent les conclusions qu'attendent les producteurs de données. L'ouverture des données publiques réduirait ainsi la fiabilité des données publiques...

## Le côté sombre des données publiques ouvertes : plausible ?



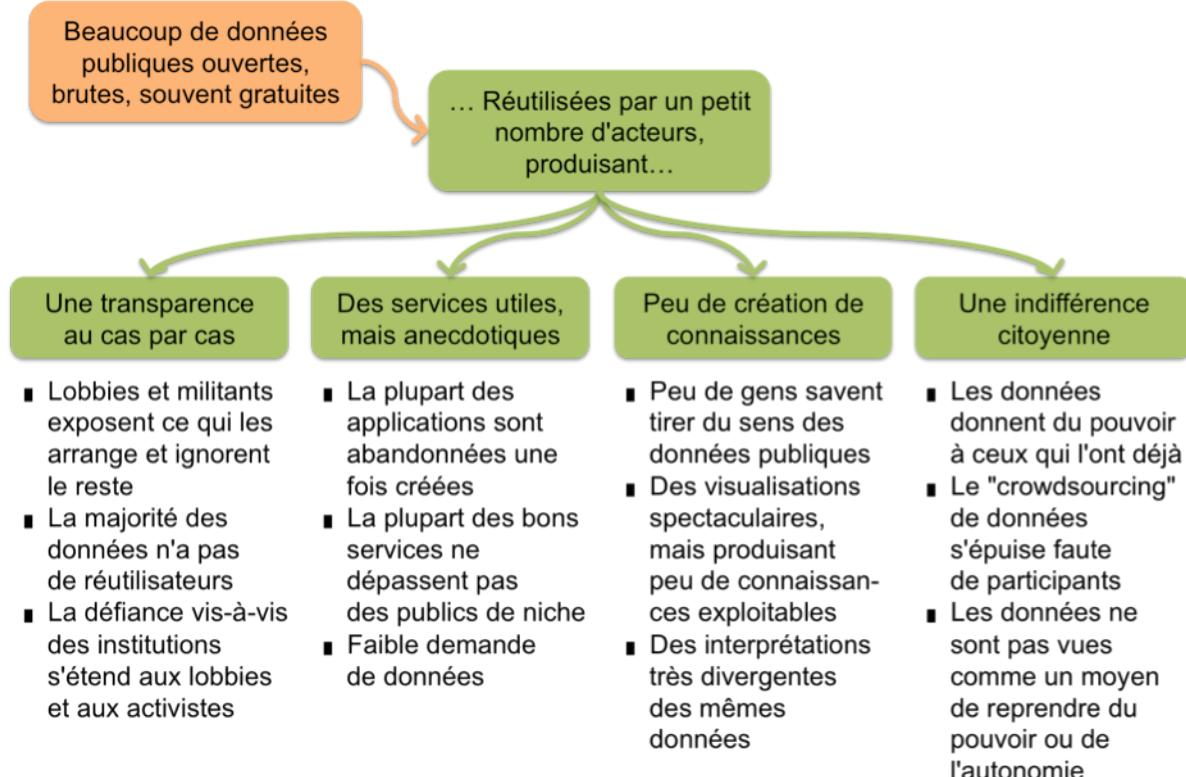
Imaginons enfin que la capacité à extraire du sens à partir des masses de données disponibles demeure très inégalement distribuée – ce qui n'apparaît pas invraisemblable. Dans un article intitulé "Du pouvoir à ceux qui l'ont déjà ?" ("Empowering the empowered?"), Mike Gurstein cite l'exemple d'une recherche sur la numérisation des registres fonciers de Bangalore (Inde) : "La mise en accès de l'information sur la propriété foncière à Bangalore a principalement permis aux personnes les plus aisées à quelques entreprises à évincer les populations pauvres et marginalisées de leurs terres." Et ici, en Europe, des lobbies bien financés ne sauront-ils pas, aussi, faire un usage plus efficace des données que de petits groupes locaux ? Les chercheurs qui voudront traiter des données ne devront-ils pas chercher leurs financements chez des acteurs qui auront intérêt à leurs résultats ? Peut-être la "sagesse des foules" corrigera-t-elle ces déséquilibres, mais rien n'est moins sûr.

## Data lassitude

Lorsque j'ai présenté ces deux premiers scénarios à l'équipe de la Fing, elle en a immédiatement produit un troisième, qui décrit une adoption modeste des données publiques, produisant des résultats anecdotiques ou décevants.

Les données publiques sont si nombreuses, complexes, hétérogènes, difficiles à interpréter, que très peu d'acteurs parviennent à réunir le courage ou les ressources pour les travailler sérieusement. Les premières applications que nous voyons aujourd'hui candidater à des concours tels qu'"Apps for Democracy" ou "Show Us a Better Way" en restent à l'état d'aujourd'hui, celui de prototypes, rarement exploitables par des citoyens ordinaires. La plupart ne se convertissent jamais en services opérationnels. De magnifiques visualisations font plus pour la célébrité de leurs concepteurs que pour créer une compréhension partagée de phénomènes complexes. Seuls quelques lobbies plongent dans les données, pour en ramener les résultats qu'ils savaient vouloir y trouver, et sans grand risque de trouver contradicteur. Après quelques contributions sur des sites comme FixMyStreet, les citoyens se fatiguent de faire le boulot qu'ils attendent de leurs municipalité. L'ouverture des données publiques ne fait pas beaucoup de mal, mais pas beaucoup de bien non plus.

## Données publiques ouvertes : l'indifférence et la déception



### Bifurcations

Chacun pourra croire qu'un de ces scénarios (ou d'autres encore) a plus de chances de se réaliser que les autres. Mais sur quelles bases ? Quelles actions, décisions, conditions, ont le plus de chances de nous engager dans une voie plutôt que l'autre ? Peut-on avoir des conséquences positives sur certains plans, négatives ailleurs, nulles ailleurs encore ?

Nous commençons à peine à envisager ces questions. Et pourtant il faut le faire, maintenant, ou courir le risque de graves déconvenues. Il y a plusieurs manières d'ouvrir les données publiques, et beaucoup de manières d'accompagner cette ouverture. Si, par exemple, on n'ouvre les données que dans l'espoir de réduire les dépenses publiques en transférant la fourniture de nombreux services au privé (voire au secteur associatif), alors la probabilité de voir s'engager les cercles vicieux décrits dans le scénario "sombre" devient très élevée.

Nous devons soulever quelques questions difficiles. Comment le rôle des acteurs publics doit-il évoluer si les entreprises et les citoyens gagnent en capacité d'action grâce aux données qu'ont produites ces mêmes institutions ? Que faudrait-il faire pour éléver la capacité de tous de lire et exploiter des données, et qui doit-il le faire ? Quelles sont les limites de la transparence : la vie privée, la sécurité, la possibilité de peser tranquillement des décisions difficiles – ou bien n'y a-t-il pas de limite ? Comment financerons-nous demain la production de données du domaine public ? Comment développer une "culture des données" qui nous rendra même capables de discuter comment les données "brutes" elles-mêmes sont produites – puisque nous savons (au moins depuis Einstein) qu'une donnée est toujours construite par l'observateur.

### Acteurs et leviers

La liste qui précède n'est pas close. Bien d'autres défis se feront jour à mesure que les Etats, les collectivités locales, les laboratoires et même des entreprises livreront leurs données, et que ces citoyens, des entreprises, des chercheurs, des activistes, des lobbyistes ou des artistes s'en empareront.

Pour commencer à identifier et à traiter ces questions, nous avons d'ailleurs intérêt à les penser à partir des acteurs : qui est concerné, qui a intérêt à la production, la circulation, la compréhension

ou l'exploitation de quelles données ? Ouvertes ou non, les données n'existent pas toutes seules, elles ne circulent pas en apesanteur, mais au contraire dans un espace parcouru de courants, de champs de force, où des intérêts, des habitudes, des capacités, coopèrent ou luttent entre eux. Personne ne tire le même pouvoir des mêmes données, dans les mêmes circonstances.

### **Qui sont les acteurs des données publiques ouvertes ?**

|   | Produc-teurs | Réutili-sateurs | Inter-médiaires | Intérêt éco. | Intérêt public |
|---|--------------|-----------------|-----------------|--------------|----------------|
| <b>Institutions publiques</b>                 |              |                 |                 |              |                |
| Nationales / Continentales ("Governments")    | ++           | +               | +               | +            | ++             |
| Locales                                       | ++           | +               | -               | +            | ++             |
| Opérateurs de services publics                | ++           | +               | -               | ++           | ++             |
| <b>Entreprises</b>                            |              |                 |                 |              |                |
| Fournisseurs spécialisés de services en ligne | +?           | ++              | ++              | ++           | -              |
| Fournisseurs de technologies                  | -            | +               | +               | +            | +              |
| Autres entreprises existantes                 | +?           | ++              | -               | ++           | -              |
| Startups                                      | -            | ++              | +               | ++           | -              |
| <b>Autres</b>                                 |              |                 |                 |              |                |
| Médias, blogueurs                             | -            | ++              | +               | +            | ++             |
| Associations, activistes, lobbyistes          | -            | ++              | +               | -            | ++             |
| Chercheurs                                    | +            | ++              | -               | ++           | ++             |
| Citoyens                                      | +            | +               | -               | +            | ++             |

Nous pouvons également penser en termes de leviers, c'est à dire des décisions, des structures et des actions qui peuvent influer sur les résultats de l'ouverture des données publiques.

Certains de ces leviers sont déjà connus, parce qu'ils émergent naturellement dès que l'on pense à ouvrir les données : les conditions d'accès (licences et tarifs), le financement des producteurs spécialisés de données, les manières de stimuler l'"écosystème" des réutilisateurs de données publiques... Tout le monde organise des concours afin de stimuler l'inventivité des entrepreneurs et développeurs, mais comment aller chercher d'autres groupes, moins motivés *a priori* ? Faut-il organiser des ateliers ou des sessions de formation à destination d'associations, de PME non-numériques, de médias locaux ? Faut-il aider certains réutilisateurs plutôt que d'autres ? Faut-il, comme le propose Tim Davies<sup>73</sup>, travailler sur des interfaces communes qui facilitent l'accès aux données de la part de non-spécialistes, sans trop influer sur l'interprétation possible de ces données ?

D'autres leviers ont un caractère plus politique, plus large. Comment diffuser une "culture des données" et éviter que la fracture sociale ne passe par la capacité de comprendre les données ? Cette "culture" doit-elle inclure une réflexion sur le bon niveau de transparence à obtenir des acteurs publics (et privés) ? Et d'ailleurs, certaines données produites par les acteurs privés doivent-elles devenir publiques ? – je pense à Adam Greenfield, s'étonnant que les données que produisent des capteurs privés installés dans l'espace public demeurent entièrement invisibles. Quel est, aussi, le statut des données produites par les citoyens ? En définitive, verrons-nous émerger une fonction de régulation du paysage des données ouvertes ?

<sup>73</sup> <http://practicalparticipation.co.uk/odi/report/>

A nouveau, la liste n'est pas close. Cet article a précisément pour but de l'ouvrir. Commençons dès aujourd'hui à explorer ensemble ce qu'il se passe *après* l'ouverture des données publiques. Pour pouvoir, dans quelques années, nous montrer fiers d'avoir contribué à engager ce mouvement.

## A2. PETIT LEXIQUE DE L'OPENDATA

---

### **API (Application Programming Interface)**

Désigne une interface de programmation fournie par un programme informatique pour permettre à d'autres programmeurs d'en utiliser certaines fonctions ou données. A travers le développement d'API, il s'agit pour un programmeur de développer une fois pour toute une application contenant tous les outils de base nécessaires, et de la mettre à disposition de la communauté des développeurs pour leur faciliter le travail.

Par exemple, les API de Google permettent à d'autres développeurs d'utiliser simplement, au sein de leur propre application ou service, des fonctions telles que la recherche ou les cartes de Google Maps. C'est par l'intermédiaire d'API (en l'occurrence, des requêtes plus ou moins préformâtes) que des sites extérieurs peuvent obtenir des données actualisées disponibles sur des sites dont ils ne sont pas les éditeurs, par exemple pour interroger une grande base de données ou récupérer des informations temps réel sur les transports.

Les API peuvent être mises à disposition de manière ouverte (accessible à tous) ou réservée à quelques acteurs, gratuite ou payante.

### **Crowdsourcing (enrichissement, approvisionnement par la "foule")**

Désigne le fait d'utiliser l'intelligence, la créativité, le savoir-faire d'un grand nombre de personnes (en l'occurrence des internautes) pour enrichir, alimenter un service, une application, un site web, la conception d'un produit, etc. L'idée sous-jacente est que la participation d'un grand groupe de personnes peut produire des connaissances, des contenus, une "œuvre collective", dont la valeur dépasse de loin celle que pourrait produire n'importe quel participant individuel.

### **Données publiques**

(voir partie 1-1)

### **Data journalism (Journalisme de données)**

Désigne une nouvelle façon de faire de l'investigation journalistique en se basant sur l'analyse de données et (souvent) la représentation visuelle. Le journaliste utilise des bases de données comme sources et en déduit des connaissances, des corrélations ou des intuitions qui ne seraient pas accessibles par les méthodes traditionnelles de l'enquête journalistique.

Même si l'article reste la composante de base, l'illustration des idées par l'image graphique, un schéma, une carte, etc., prend une place plus importante

### **Linked Data (les données liées)**

Désigne une approche du web poussée par les tenants du "web sémantique" qui encourage à décrire toutes les données d'une manière lisible par des ordinateurs, et à les lier entre elles en décrivant leurs relations, ou encore, en facilitant leur mise en relation : ainsi, deux informations produites par des acteurs différents, mais correctement décrites comme faisant référence à un même territoire, pourront être "liées" même si leurs producteurs n'y auraient jamais pensé. Les Linked Data sont liées les unes avec les autres à l'aide de certains "formats" (propres au web sémantique : RDF, OWL, SIOC...) lisibles par un ordinateur. Celui-ci devient alors capable de les associer de manière automatique. Tim Berners-Lee, "père du web" et champion du Linked Data, a contribué à la mise en place du portail gouvernemental britannique des données publiques, [data.gov.uk](http://data.gov.uk).

### **Interopérabilité**

Désigne la capacité que possède un produit ou un système, dont les interfaces ont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre<sup>74</sup>.

---

<sup>74</sup> Définition donnée par Wikipédia <http://fr.wikipedia.org/wiki/Interopérabilité>

## **Métadonnée**

Désigne un ensemble structuré d'informations servant à définir ou décrire une ressource (donnée, document...) quel que soit son support (papier ou électronique). S'agissant d'un livre par exemple, le titre, le nom d'auteur, l'éditeur, l'année de publication en constituent des métadonnées.

Les métadonnées sont donc "des données sur les données", qui permettent à un individu ou un ordinateur d'en comprendre le sens et l'organisation. Elles sont à la base des techniques du web sémantique<sup>75</sup>.

## **Mashup (combiner, mélanger ensemble)**

Désigne une pratique consistant à combiner différentes sources (le plus souvent externes) d'informations ou de données afin d'alimenter son propre site web. De plus en plus de sites web (Google, Yahoo!, Amazon, eBay...) proposent gratuitement des API, afin d'encourager la communauté des développeurs à créer des mashup utilisant leur contenu.

## **Web 2.0**

Il n'existe pas de définition stable. L'expression a été proposée en 2004 par l'éditeur Tim O'Reilly en 2004 pour décrire une certaine évolution du web fondée sur quatre piliers :

- Le recentrage de l'information et des services autour de la personne ;
- Une facilitation considérable de l'intervention des utilisateurs dans la production et l'échange de textes, images et vidéos ;
- Des interfaces standards permettant de partager et d'agencer de diverses manières informations et services ;
- Et une expérience plus "sociale" du web par laquelle l'innovation et la valeur proviennent, dans une large mesure, de l'interaction entre les utilisateurs.

Les services et plateformes emblématiques du web2.0 sont Wikipedia, eBay et Craigslist, Youtube, Facebook et Twitter.

## **Web<sup>2</sup> ("web squared", le web puissance 2) :**

L'expression est proposée par Tim O'Reilly et John Battelle, pour décrire le développement non plus arithmétique mais exponentiel des possibilités du web. Le web "1.0" reliait des documents, le web "2.0" reliait des gens, le web au carré relie des informations individuelles, des données. Pour certains, cette expression un peu *ad hoc* désigne une phase intermédiaire avant le "web sémantique", pour d'autres, les deux expressions désignent un peu la même chose.

## **Web sémantique ou web des données :**

Désigne un ensemble de technologies visant à rendre l'ensemble des ressources du web (documents et "grains" élémentaires d'informations) accessible,, intelligibles et utilisables par des programmes et agents logiciels, grâce à un système de métadonnées. Les machines pourront ainsi traiter, relier et combiner automatiquement un certain nombre de données.

---

<sup>75</sup> Définition adaptée de celle de Wikipédia <http://fr.wikipedia.org/wiki/Métadonnée>

## A3. REPERTOIRE DES ACTEURS, ORGANISMES ET INSTITUTIONS

Cette sélection, qui ne prétend pas à l'exclusivité, recense les organismes et groupements actifs dans le domaine de l'ouverture des données publics.

### En France

|  |   |
|--|---|
| <u><a href="#">APIE</a></u> – Agence du patrimoine immatériel de l'Etat        | <p>L'agence, créée en septembre 2007, est un service à compétence nationale, rattaché conjointement au directeur général du Trésor et de la politique économique et au directeur général des Finances publiques.</p> <p>Ses missions sont :</p> <ul style="list-style-type: none"><li>• De sensibiliser les acteurs publics aux enjeux attachés aux actifs immatériels de l'État ;</li><li>• de faciliter la prise en compte de ces enjeux et l'adoption de nouveaux modes de gestion au sein de l'administration ;</li><li>• d'accompagner les administrations et les établissements publics pour la définition de nouvelles stratégies et la conduite de projets concrets de valorisation de leur patrimoine immatériel.</li></ul> <p>Depuis sa création, l'APIE mène des actions concrètes pour aider les opérateurs économiques à mieux accéder aux informations publiques réutilisables et pour sensibiliser et accompagner les administrations dans la mise à disposition de leurs informations. L'objectif est de favoriser l'émergence d'un cadre pratique, juridique et financier homogène de nature à faciliter la réutilisation.</p> |
| <u><a href="#">Cada</a></u> - Commission d'accès aux documents administratifs  | <p>La CADA est une autorité administrative indépendante dont le rôle est consultatif. Elle est chargée de veiller au respect de la liberté d'accès aux documents administratifs et aux archives publiques. Elle émet des avis lorsqu'elle est saisie par une personne qui rencontre des difficultés pour obtenir la communication d'un document administratif ou pour consulter des documents d'archives publiques. La saisine de la CADA pour avis est un préalable obligatoire à l'exercice d'un recours contentieux.</p>   |
| <u><a href="#">DGME</a></u> – Direction générale de la modernisation de l'Etat | <p>Créée en décembre 2005, la DGME est une direction du ministère du Budget, des comptes publics, de la fonction publique et de la réforme de l'Etat. Elle capitalise l'expérience acquise dans la conduite de nombreux projets de transformation et conseille les ministères dans leurs stratégies de transformation. Elle les accompagne dans la mise en œuvre des plans de transformation et des décisions adoptées dans le cadre de la Révision générale des politiques publiques (RGPP), lancée en juin 2007 par le Gouvernement.</p> <p>L'enjeu : transformer l'État pour maîtriser et rationaliser les dépenses publiques tout en améliorant la qualité des politiques publiques et le service rendu aux usagers.</p> <p>A l'écoute des usagers et de leurs attentes, elle identifie les priorités et élabore la stratégie d'amélioration des services rendus aux usagers et pilote plusieurs chantiers interministériels structurants dans les domaines de la simplification, de l'accueil et de l'administration électronique.</p>   |

|  |   |
|--|---|
| <a href="#">Fing</a>   | <p>La Fing est une association dont la vocation est de produire et partager des idées neuves pour anticiper les transformations numériques. Elle aide les grandes entreprises et les start-ups, les territoires et les décideurs politiques, les chercheurs, les créateurs, les innovateurs sociaux... à anticiper les opportunités et les risques associés aux technologies, à leurs usages et au système d'innovation qui les accompagne.</p> <p>La Fing détecte les signaux faibles et décrypte les tendances lourdes. Elle repère, met en valeur et en relation les projets les plus innovants. Elle identifie de nouvelles opportunités, de nouvelles pistes d'innovation. Elle agit en expérimentant des projets transformateurs. L'ouverture des données publiques est l'un de ces projets.</p>  |
| <a href="#">GFII</a> - Groupement français de l'industrie de l'information | <p>Le GFII est une association créée en 1979 réunissant les principaux acteurs publics et privés de l'industrie de l'information et de la connaissance en France.</p> <p>Un de ses groupes de travail réunit les experts de la réutilisation professionnelle des données publiques. Ses missions : inciter et aider les producteurs publics à favoriser la diffusion de leurs données dans des conditions économiques viables pour toutes les parties en présence.</p>  |
| <a href="#">Regards citoyens</a>   | <p>Le collectif RegardsCitoyens.org est une association constituée de citoyens de tous âges et régions mus par un désir commun de proposer un accès simplifié au fonctionnement de nos institutions démocratiques.</p> <p>Regards Citoyens a notamment créé le portail "<a href="#">Nos Données</a>", qui vise à apporter plus de visibilité aux données publiques librement accessibles à chacun</p>   |
| <a href="#">AEC</a> – Aquitaine Europe Communication                       | <p>AEC accompagne les pouvoirs publics dans la définition et la mise en œuvre de leurs stratégies numériques, notamment par :</p> <ul style="list-style-type: none"> <li>• l'appui à la concertation et la production d'outils d'aide à la décision,</li> <li>• la contribution à l'émergence et l'accompagnement de projets d'intérêt public partenariaux, d'entreprises et de dispositifs innovants dans le numérique,</li> <li>• la diffusion des cultures du numérique.</li> </ul> <p>AEC se positionne comme tiers de confiance, médiateur d'informations qualifiées, fournisseur d'analyses stratégiques, coproducteur de dispositifs innovants et développe une attitude de proximité dans l'élaboration et l'accompagnement de projets.</p> <p>AEC est l'auteur d'un "<a href="#">Guide pratique et juridique</a>" sur les données publiques.</p> |
| <a href="#">LiberTIC</a>   | <p>LiberTIC est une association nantaise créée en décembre 2009 suite au premier <a href="#">Open Data Camp</a> de Paris. Elle a pour objectif de promouvoir l'ouverture des données publiques, l'e-démocratie, le gouvernement 2.0 et d'accompagner le territoire ligérien dans le développement et l'utilisation d'outils innovants.</p> <p>Parallèlement, Libertic participe au mouvement de rapprochement de l'Economie sociale et solidaire et des TIC afin de créer des synergies au niveau local.</p>  |

## En Europe

|  |  |
|--|--|
| <a href="#"><u>PSI Alliance</u></a>              | PSI Alliance est une association qui a été créée en 2008 pour encourager le secteur public à s'engager résolument dans la diffusion de ses données. L'association, structurée au niveau européen, regroupe à ce jour essentiellement des grandes entreprises réutilisatrices de données publiques. Véritable réseau d'influence transfrontalier, et sensibilisant les parlementaires européens, elle s'attache à participer au développement d'une industrie numérique européenne profitant à la fois aux secteurs public et privé et aux citoyens.  |
| <a href="#"><u>Open Knowledge Foundation</u></a> | L'Open Knowledge Foundation est une organisation à but non lucratif fondée en 2004, qui promeut l'ouverture de la connaissance et du savoir sous toutes ses formes.<br><br>L'organisation a mis en place un annuaire des données publiques mondial, le CKAN (Comprehensive Knowledge Archive Network), alimenté essentiellement par ses utilisateurs. Début 2011, l'annuaire contenait 1557 "data packages" (dont 8 répondent au mot clé "France").  |
| <a href="#"><u>ePSI Platform</u></a>             | La Plateforme européenne de l'information de service public est un carrefour d'informations et d'acteurs qui se fixe pour objectif de révéler le potentiel de valeur de l'information publique à travers l'Europe.<br><br>Elle se constitue d'une petite équipe qui travaille à la croisée de la communauté des acteurs réutilisateurs de données et des acteurs publics participant à l'ouverture des données. L'action de l'ePSI Platform vise à stimuler les actions de réutilisation des données publiques d'un point de vue national et européen, et à participer au développement d'un environnement plus solide et plus transparent dans le champ de l'information publique. Elle anime un site web qui communique sur l'actualité européenne (événements, rapports, actualité législative, etc.). L'ePSI Platform est soutenue par la Commission européenne dans le cadre de son programme eContentplus. |

## A l'international

|   |   |
|---|---|
| <a href="#"><u>Open Data Foundation</u></a><br>(ODaF) | Cette association américaine a pour objectif d'améliorer l'accessibilité des données et métadonnées pour une meilleure circulation de l'information scientifique, publique, économique, etc.<br><br>Elle travaille pour cela à la définition et l'adoption de métadonnées standards permettant le développement de solutions open source pour l'utilisation des données statistiques. |
|---|---|

## A4. BIBLIOGRAPHIE - WEBOGRAPHIE

---

### Rapports - Livre blancs

Rapport "Partager notre patrimoine culturel", sous la direction de Bruno Ory-Lavollée, remis au ministre de la Culture en juillet 2009 : "How is open government data being used in practice", Tim Davies, Practical Participation : <http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf>

Livre blanc "Anonymisation des Données à caractère personnel", Bruno Rasle : <http://www.cortina.fr/livres-blancs.php>

"Open Data Open Society", Marco Fioretti, Scuola Superiore Sant'Anna, Pisa, rapport dans le cadre du réseau d'excellence européen DIME (Dynamics of Institutions and Markets in Europe), 2010 : <http://www.lem.sssup.it/WPLem/odos/odos.html>

### Articles

"Mesurer les effets de la libération de données", Hubert Guillaud, InternetActu : [www.internetactu.net/2009/09/16/critiques-du-web2-24-les-effets-de-la-liberation-des-donnees/](http://www.internetactu.net/2009/09/16/critiques-du-web2-24-les-effets-de-la-liberation-des-donnees/)

"Les joyaux de la couronne n'appartiennent à personne", Jean-Marc Manach, InternetActu : [www.internetactu.net/2008/12/03/les-joyaux-de-la-couronne-nappartiennent-a-personne/](http://www.internetactu.net/2008/12/03/les-joyaux-de-la-couronne-nappartiennent-a-personne/)

"Du contenu roi aux données reines", Fred Cavazza : [www.fredcavazza.net/2010/07/19/du-contenu-roi-aux-donnees-reines/](http://www.fredcavazza.net/2010/07/19/du-contenu-roi-aux-donnees-reines/)

"Les enjeux publics économiques et citoyens de l'ouverture des données : l'expérience britannique", Valérie Peugeot, Orange Labs, 2010 : [http://laborange.academia.edu/Val%C3%A9riePeugeot/Papers/343143/Les\\_enjeux\\_publics\\_economiques\\_et\\_citoyens\\_de\\_louverture\\_des\\_donnees\\_lexperience\\_britannique](http://laborange.academia.edu/Val%C3%A9riePeugeot/Papers/343143/Les_enjeux_publics_economiques_et_citoyens_de_louverture_des_donnees_lexperience_britannique)

"Web des données, données ouvertes", Duprat J., Tramblay D., Peugeot V., Orange Labs, 2009 : [www.orange-innovation.tv/dreamorange/resources/documents/4022/OpenData\\_OrangeLabs\\_Sofrecom\\_2009.pdf](http://www.orange-innovation.tv/dreamorange/resources/documents/4022/OpenData_OrangeLabs_Sofrecom_2009.pdf)

"Les bons ingrédients pour une ouverture des données réussie", Caroline Goulard, OWNI, 2010 : <http://owni.fr/2010/05/31/opendata-12-data-gov-ou-data-gov-uk/>

"Des licences libres pour concilier innovation sociale et économique", Regards Citoyens, 2010 : <http://owni.fr/2010/06/14/des-licences-libres-pour-concilier-innovation-sociale-et-economique/>

"Les données climatiques devraient devenir publiques", Gilles Mahé et Alain Gioda, *Le Monde*, 2009 : [www.lemonde.fr/idees/article/2009/12/21/les-donnees-climatiques-devraient-devenir-publiques-par-gil-mahe-et-alain-gioda\\_1283739\\_3232.html](http://www.lemonde.fr/idees/article/2009/12/21/les-donnees-climatiques-devraient-devenir-publiques-par-gil-mahe-et-alain-gioda_1283739_3232.html)

### Webographie

La webographie sur ce sujet est abondante et en constante évolution. Nous renvoyons donc à la webographie collaborative publiée sur le réseau social de la Fing : <http://www.reseaufing.org/pg/pages/view/16823/>