HAL
open science

L'Open Data au service de la connaissance territoriale ? Le développement de la démarche Open Data de la Région Nord-Pas De Calais

Louis Platt

To cite this version:

Louis Platt. L'Open Data au service de la connaissance territoriale ? Le développement de la démarche Open Data de la Région Nord-Pas De Calais. Environnement et Société. 2015. dumas-01274336

HAL Id: dumas-01274336

https://dumas.ccsd.cnrs.fr/dumas-01274336

Submitted on 15 Feb 2016

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

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

UNIVERSITÉ DE GRENOBLE

Université Joseph Fourier Grenoble

Institut de Géographie Alpine

Cité des Territoires

Rapport de fin d'étude de Master 2

IDT

Ingénierie du Développement Territorial

L'Open Data au service de la connaissance territoriale ?

Le développement de la démarche Open Data de la Région Nord-Pas De Calais

présenté par Louis PLATT

platt.louis@gmail.com

OPEN DATA Région Nord-Pas de Calais

Le 16/06/2015 à l'Institut de Géographie Alpine, Grenoble

Stage effectué au Conseil Régional Nord-Pas de Calais
151, Avenue du Président Hoover, 58000 Lille
sous la direction de Joël TIGNON et Adrien CARPENTIER
Tuteur pédagogique : Grégoire FEYT

NOTICE ANALYTIQUE Rapport de fin d'étude

M2 IDT –Ingénierie du Développement Territorial

PARCOURS : ☐ IDT-stratégie ☐ IDT-animation ☒ IDT-géomatique

ANNEE UNIVERSITAIRE : 2014/2015

OBSERVATIONS DU JURY

A : Très bon rapport de fin d'étude
B : Bon rapport mais avec des faiblesses sur quelques aspects
C : Rapport acceptable présentant un intérêt particulier sur certains aspects

Précisions éventuelles sur les faiblesses et les forces du mémoire :

Signatures

AUTEUR	NOM		PRÉNOM	
	PLATT		Louis	
TITRE	L'Open Data au service de la connaissance territoriale ? Le développement de la démarche Open Data de la Région Nord-Pas De Calais			
UNIVERSITÉ JOSEPH FOURIER & Institut de Géographie Alpine	Nom et prénom du Tuteur pédagogique	Nom de la structure dans laquelle le stage a eu lieu (et localisation)	Nom et prénom du responsable professionnel présent dans le jury	
	Grégoire FEYT	Conseil Régional Nord-Pas de Calais 151, Avenue du Président Hoover, 58000 Lille	**Adrien CARPENTIER**	
COLLATION	Nb. de pages	Nb. de volumes	Nb. d'annexes	Nb. de réf. biblio.
	77	1	7	45
MOTS-CLÉS	Open Data, Connaissance Territoriale, Collectivités territoriales, Animation			
TERRAIN D'ÉTUDE OU D'APPLICATION	**Région Nord-Pas De Calais**			

Les démarches Open Data fleurissent en France ces dernières années L'ouverture large et importante des données des collectivités territoriale constitue un potentiel « d'objectivation » nouveau, et participerait à l'amélioration de la connaissance territoriale, elle-même nécessaire à la prise de décision et au développement territorial Si ces dispositifs mettent à disposition de nombreuses nouvelles données, ils nécessitent également d'être accompagnés d'une animation importante et ce afin de mettre le citoyen au cœur du développement territorial.

The French Open Data initiative increase their number those last years. A large and massive open data from local authorities constitute a new potential for local knowledge which is needed for decision making and local development. If those operations provide a lot of new datas, it also need specific animation to allow citizen to take their place in local development

Remerciements :

Je tiens à remercier Grégoire Feyt pour ses conseils et sa disponibilité,

Je remercie également Joël Tignon et Adrien Carpentier pour leur suivi, et leur accompagnement tout au long de mon stage.

Je remercie également Florent Lamiot et Adrien Carpentier pour m'avoir initié au monde du libre et du collaboratif.

Enfin je remercie l'ensemble de l'équipe pédagogique du master IDT de m'avoir permis de développer des connaissances et des compétences que je saurai mettre à contribution plus tard.

Et bien sûr merci aux étudiants d'IDT.

Table des matières

Introduction : 7

I. Qu'est-ce que l'Open Data : 8

A. L'apparition de l'Open Data : 8

1. L'Open Data et la science : 8

2. Le cas de l'Information géographique : 9

3. L'Open Data et la « culture libre » : 10

4. L'Open Data comme enjeu politique : 10

B. L'Open Data en France : 11

1. La démarche Open Data… : 11

2. … utilise la donnée publique …: 12

3. …soumise à des licences … : 13

4. …et s'inscrit dans une stratégie plus large : 15

Conclusion de la première partie : 17

II. Open Data et connaissance territoriale : 18

A. La donnée comme outil rationnel d'aide à la décision : 18

1. L'ouverture des données référentielles pour faire sens commun: 18

2. L'ouverture des données métiers, et création de connaissance territoriale : 19

3. L'interopérabilité des données ouvertes : 23

4. Les acteurs, créateur de connaissance territoriale : 24

B. Vers une meilleure mobilisation des citoyens au service de la connaissance territoriale : L'animation de la démarche Open Data. 26

1. Mobiliser les citoyens « éclairés » pour créer une communauté : 27

2. Rendre le citoyen acteur de la connaissance territoriale : 29

3. Les acteurs de la médiation : 30

Conclusion de la deuxième partie : 33

III. La Stratégie Open Data de la Région Nord-Pas de Calais : 34

A. Mise en contexte du stage : 34

1. Présentation de la structure : 34

2. Eléments de contexte : 35

3. Commande et réponses proposées : 37

B. Ma mission au sein de la Région Nord-Pas De Calais révélatrice des enjeux et limites des dispositifs Open Data : 41

1. L'Amélioration de la plateforme : 41

2. L'enrichissement de la plateforme : 42

3. Animation de la démarche : 46

4. La réutilisation des données : 48

C. Retour d'expérience : 48

Conclusion de la troisième partie : 49

Conclusion : 50

Annexes: 52

Annexe 1: Open Government Data Principles 52

Annexe 2 : Offre de stage 53

Annexe 3 : méthodologie de Benchmark 55

Annexe 4 : Tableau des jeux de données repérés comme ouvrable au sein du serveur du service IGAS. 63

Annexe 5 : Présentation Open Data 64

Annexe 6 : fiche de présentation Open Data 70

Annexes 7 : Note de synthèse donnée de santé 71

Table des figures : 73

Table des sigles : 74

Bibliographie : 75

Introduction :

Depuis 2010, les démarches Open Data sont apparues en France. Au départ, portées par quelques collectivités pionnières (Rennes, Paris…), les initiatives se sont multipliées et bénéficient aujourd'hui d'une réelle volonté politique tant à l'échelle nationale, qu'à l'échelle locale. Aux côtés de ces démarches institutionnelles relativement récentes, des initiatives plus anciennes, issues du monde du « libre » ou du collaboratif (OpenStreetMap, wikimédias) participent également à l'explosion des données ouvertes disponibles.

L'ouverture des données (Open Data) constitue une source d'information extrêmement importante, qu'il est possible de mobiliser au service du développement territorial. Selon Feyt et Lardon, « la décision territoriale est indissociable de la notion de connaissance ». Cette connaissance renvoyant à deux registres d'expertises différents : l'un subjectif basé sur une approche sensible et l'autre objectif, basé sur des éléments quantifiables ou qualifiables (Feyt, Lardon 2010). Ainsi le développement territorial se base sur la connaissance territoriale. Mais qu'entendons-nous par connaissance territoriale ?

La connaissance territoriale découle d'un processus de « raffinage » et renvoi à trois registres distincts (Feyt, 2007). Le premier registre sur lequel se base la connaissance territoriale est celui de l'information territoriale, cette information étant elle-même issue de l'interprétation, du croisement de données relatives au territoire, à ses habitants ou aux interactions y ayant lieu. On a donc ici un premier niveau de raffinage, on part de la donnée brute qui est un élément ambigu en « attente »Chignard, 2012)) auquel on applique un traitement, une interprétation pour faire apparaître de l'information. C'est le registre de « l'objectivation technique ». Le second registre de la connaissance territoriale est le registre de la compréhension des phénomènes et des dynamiques en œuvre au sein du territoire, impliquant ici la question de l'observation et de l'analyse. Ce registre nécessite de se placer selon une posture, et tient donc d'une part de subjectivité. Enfin le dernier registre est celui de l'utilisation de cette connaissance, ce qui renvoie aux choix politiques, et aux orientations souhaitées pour le développement du territoire. On parle de connaissance territoriale, ce qui implique qu'elle repose sur des systèmes d'acteurs, différentes échelles (temporelles et spatiales) et des représentations.

L'ouverture large et importante des données constitue donc un potentiel « d'objectivation », et participerait à l'amélioration de la connaissance territoriale, elle-même nécessaire à la décision et donc au développement territorial. Dès lors il paraît intéressant de s'interroger sur l'impact de l'Open Data sur la création et le partage de la connaissance territoriale.

Afin de répondre à cette question en nous appuyant sur les différentes études et documents scientifiques ou institutionnels produits ainsi que sur notre expérience de stagiaire au sein du Conseil Régional Nord-Pas de Calais, nous nous intéresserons dans un premier temps à ce qu'est aujourd'hui l'Open Data. Dans un second temps nous verrons comment l'ouverture des données participe à la création de connaissances territoriales. Dans un dernier temps en nous appuyant sur l'exemple du développement de la stratégie Open Data de la Région Nord-Pas De Calais, nous chercherons à mettre en évidence les freins et les limites de ces démarches Open Data.

I. Qu'est-ce que l'Open Data :

A. L'apparition de l'Open Data :

1. L'Open Data et la science :

L'Open data tire sa source du monde scientifique. Dès 1942 le sociologue Robert King Merton dans *the normative structure of science* explique que pour faire avancer la connaissance, les chercheurs doivent rendre accessible librement et à tous les résultats de leurs travaux et donc renoncer à leurs droits de propriété intellectuelle pour contribuer à une avancée générale de la science (Merton 1942). C'est l'apparition de l'idée de connaissance comme partie intégrante des biens communs. Cette théorie sera reprise, retravailler et élargit pour toucher d'autres thèmes que la connaissance, des années plus tard avec l'apparition et la libéralisation des Technologies de l'Information et de la Communication(TIC). Ainsi des chercheurs comme la prix Nobel d'économie 2009 Elinor Olstrom ou l'Informaticien Philippe Aigrain feront naître le concept de biens communs informationnels (Aigrain 2005) dont nous garderons la définition donnée ci-dessous.

« BIENS COMMUNS. Toute « chose » ou entité immatérielle à laquelle on a décidé de donner un statut de propriété commune, de la faire appartenir à tous, parce qu'elle n'appartient à personne. Dans le sens moderne, la propriété commune est universelle, elle est celle de l'humanité. Dans le sens ancien, il s'agissait souvent de la propriété d'une communauté restreinte. À ne pas confondre avec les biens publics dans le sens d'objets d'une propriété publique (gérée par des institutions publiques) ». (Aigrain 2005, p265)
« BIENS COMMUNS INFORMATIONNELS. Biens communs qui peuvent être créés, échangés et manipulés sous forme d'information, et dont les outils de création et le traitement sont souvent eux-mêmes informationnels (logiciels). Il peut s'agir de données, de connaissances, de créations dans tous les médias, d'idées, de logiciels. Les biens communs informationnels sont des biens publics parfaits au sens économique, contrairement aux biens communs physiques, qui gardent toujours une part de rivalité ou d'excluabilité. » (Aigrain 2005, p266))

Outre ce besoin de mutualiser les informations en matière de recherche, qui est apparu très tôt parmi les scientifiques. La première mention du terme « Open Data » provient également de ce milieu. Il apparait en 1995 dans un rapport du Conseil national de la Recherche aux États-Unis, intitulée *"De l'échange complet et ouvert des données scientifiques*". Ce rapport traite du nécessaire besoin d'ouvrir et de mutualiser les données géophysiques et environnementales, afin de comprendre et d'analyser les phénomènes globaux qui leur sont liés.

Enfin en 2001, lors d'un congrès international à Budapest est lancé l'Open Acces Initiative (OAI), qui vise à favoriser l'échange et la diffusion des archives numériques et des publications scientifiques de façon ouvertes (Budapest Open Access Initiative, 2001). En France ce mouvement se traduira par la mise en place en 2005 du portail commun de diffusion HAL (et ses sous domaines HALSHS par exemple). Des organismes de recherche français comme le Centre National de Recherche Scientifique (CNRS), l'Institut National de la Recherche Agronomique (INRA) permettent et poussent leur chercheur à diffuser leurs recherches sur cette plateforme.

2. Le cas de l'Information géographique :

Historiquement la maîtrise de la création et de la diffusion de l'information géographique a été rattachée à l'État Ainsi des organismes comme l'Institut nationale de l'information géographique forestière (IGN) ou l'institue Nationale de la Statistique et des Études Économiques (INSEE) produisaient les données de références. Avec la diffusion des outils informatiques les collectivités locales ont pu se dotées pour certaines de Système d'Information Géographique(SIG), créant et compilant encore plus d'informations géographiques sur leur territoire. Ces organismes avaient également pour mission d'établir des cadres, des référentiels précis pour ces données.

Si jusqu'au milieu des années 90 l'Information géographique était l'affaire de spécialistes, des échanges de données avaient lieu entre les différents producteurs ou SIG locaux, par convention (payante ou gratuite). Durant cette période les sigistes ont également beaucoup travaillé sur la question de la normalisation de leurs données, en effet afin que les données des différents SIG puissent être utilisées ensemble, les services techniques ont ouvert le dialogue, se sont regroupés. En 1986 naissait l'Association Française pour l'information géographique(AFIgéo), qui œuvre depuis pour le développement du secteur de l'information géographique. De ces discussions sont nés un certain nombre de méthodes, de principes visant à standardiser les données, les usages des SIG afin de proposer une donnée de qualité, si possible interopérable. Grâce à ce travail, les SIG sont devenus des structures reconnus au service de la connaissance territoriale.

Depuis une dizaine d'années, ce système a changé. Les collectivités locales ont vu leurs compétences s'étendre, suite aux différents actes de la décentralisation. Les limites administratives s'estompent, les territoires se composent et se recomposent au gré des projets, faisant naître de nouveaux besoins en matière de connaissance territoriale. Ces changements obligent les SIG des différentes collectivités à travailler ensemble, à mutualiser leurs données. C'est dans cette logique qu'ont été créés les observatoires territoriaux

A partir des années 2000, l'information géographique qui était jusqu'alors affaire de techniciens, de professionnels (Feyt 2011) se libéralise et s'ouvre au grand public. Les nouvelles exigences liées à internet (instantanéité, universalité…) favorisent l'émergence de demande spécifique et de qualité pour l'accès à des cartes thématiques auxquelles les instituts nationaux (dont le modèle économique était basé sur la vente des données dont ils étaient créateurs et propriétaires) ont du mal à répondre. La création et la large utilisation des systèmes de géolocalisation type GPS amène à la création de nouvelles données géographiques, « *parfois totalement hétérodoxes aux yeux d'un géomètre ou d'un géomaticien mais parfaitement efficaces au regard des usages envisagés.* »(Feyt 2011 p5). En outre l'hégémonie des instituts nationaux en matière de données de références est mise à mal, en effet le géant du web Google, après avoir racheté les fonds cartographiques de l'IGN, a développé et commercialisé Google maps, une application de cartographie en ligne aujourd'hui largement utilisée.

L'avènement d'internet a également fait apparaître de nouvelles initiatives en matière de SIG, ainsi en 2004, le projet de cartographie collaboratif OpenStreetMap est lancé, ce projet vise à proposer une base de donnée géographique libre et la plus complète du monde. Avec plus d'un million de contributeurs, cette carte fait figure de fer de lance d'une nouvelle forme d'information géographique, qui met l'utilisateur, le citoyen au cœur de la création de l'information.

3. L'Open Data et la « culture libre » :

La « culture libre » est un mouvement social apparu à la fin des années 80, ce mouvement tire sa source de l'informatique et des projets de logiciels libres. Le mouvement du logiciel libre initié par Robert Stallman en 1983, vise à proposer des logiciels libres (free software) basés sur quatre libertés essentielles *:*

la liberté d'exécuter le programme comme vous voulez, pour n'importe quel usage (liberté 0); la liberté d'étudier le fonctionnement du programme, et de le modifier pour qu'il effectue vos tâches informatiques comme vous le souhaitez (liberté 1) ; l'accès au code source est une condition nécessaire ; la liberté de redistribuer des copies, donc d'aider votre voisin (liberté 2) ; la liberté de distribuer aux autres des copies de vos versions modifiées (liberté 3) ; en faisant cela, vous donnez à toute la communauté une possibilité de profiter de vos changements.(FreeSoftware Foundation,2014)

Ainsi le logiciel libre se caractérise par un libre accès au code source du logiciel, permettant un réel contrôle du logiciel et de l'ordinateur par l'utilisateur(à la différence des logiciels propriétaire qui « bride » et oriente l'utilisation de la machine), la montée en qualité du logiciel par ses modifications par une communauté, en effet pour Stallman, plus la communauté de ré utilisateur est grande, plus les erreurs, les dysfonctionnements et les suggestions d'amélioration apparaissent et remontent rapidement. Enfin ces logiciels libres, sont basés sur l'utilisation de formats ouverts afin de favoriser l'interopérabilité, et édités sous des licences libres. Stallman popularise notamment le terme de Copyleft pour ces licences libres (jeu de mots) en opposition avec le terme copyright (droit d'auteur) des licences classiques

De ce mouvement est née une réelle philosophie, de nombreuses licences libres ont vu les jours, Ainsi Lawrence Lessings crée les licences ouvertes Creative commons en 2001, ces licences libres proposent de libérer les œuvres des droits de propriétés intellectuelles qui leur sont rattachés. Elles sont fondées sur l'idée de Copyleft et favorisent la diffusion de la connaissance. Les différents niveaux de licences fixent le degré d'ouverture et les règles de réutilisations de l'œuvre. (Ces licences sont notamment utilisées par wikimédias ou encore pour ce rapport de fin d'études). C'est également en 2001 qu'est lancée l'encyclopédie collaborative Wikipédia, qui compte aujourd'hui plus d'un million six cents mille articles pour sa version francophone.

En décembre 2007, trente penseurs personnalités de l'internet et du logiciel libre se sont retrouvés à l'initiative de Tim O'Reilly (l'inventeur du terme web 2.0), pour définir le concept d'Open government data et la faire adopter par les candidats à l'élection présidentielle américaine (Malamud, 2007). Ainsi huit principes de base sont définis (voir annexe 1).Ainsi, au même titre que les idées scientifiques les données publiques sont un bien commun. Les manières de faire portent principalement sur le partage et l'utilisation de ce bien commun, directement inspirées de la démarche et des pratiques de l'open source, et sont organisées autour de trois notions : ouverture, participation et collaboration.

4. L'Open Data comme enjeu politique :

En 2009, le jour de son investiture, le nouveau président américain Barack Obama signe deux mémorandums présidentiels sur la transparence et l'ouverture des données. Ces mémorandums précisent que l'action publique, tout comme les modèles de l'open source, du libre sera pensée autour du triptyque « transparence, participation et collaboration". (Obama, 2009).

La même année, la plateforme nationale américaine data.gov voit le jour, cette plateforme permet au citoyen d'avoir accès à l'ensemble des données gérées par la maison blanche et le gouvernement fédéral (hors données sensibles ou individuelles).

En Europe, C'est le Royaume-Unis qui se saisit le premier du sujet, emboîtant le pas au gouvernement américain, sous l'impulsion de l'inventeur du « world wide web » Tim Berners-Lee, alors conseiller du gouvernement sur les questions d'ouverture des données, le portail data.gov.uk est mis en place. Aujourd'hui le Royaume-Uni est considéré comme l'État étant le plus avancé en matière d'Open Data à l'échelle mondiale.

Au niveau de l'Union Européenne, il n'y a pas encore de réel politique relative à l'ouverture des données, cependant deux directives distinctes : l'une concernant la réutilisation des informations du secteur public (directive PSI) de 2003 et l'autre établissant l'infrastructure d'information géographique dans la Communauté Européenne(directive INSPIRE) de 2007, vise à établir un cadre commun afin d'harmoniser les informations et données et de favoriser leur réutilisation et l'interopérabilité des données des États de l'Union Européenne. Ces deux directives participent à la réflexion et au déploiement des stratégies Open Data en Europe et en France.

B. L'Open Data en France :

1. La démarche Open Data… :

Preuve s'il en est que l'Open Data a fait son chemin en France depuis près de cinq ans, le terme fait son entrée cette année dans les dictionnaires grand public que sont le Robert et le Larousse. La définition donnée à l'Open Data dans ce dernier est

Open data : n.m. invariable. Mise à disposition de données numériques, d'origine publique ou privée, libres d'accès et réutilisables par tous.

Pour les contributeurs de Wikipédia la donnée ouverte se définit comme :

« Une donnée ouverte est une donnée numérique d'origine publique ou privée. Elle peut être notamment produite par une collectivité ou un service public (éventuellement délégué) ou une entreprise ou un particulier qui la met volontairement à disposition de tous
Elle est diffusée de manière structurée selon une méthodologie et une licence libre garantissant son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière (ou avec très peu de restrictions).
L'ouverture des données (en anglais open data) représente à la fois un mouvement, une philosophie d'accès à l'information et une pratique de publication de données librement accessibles et exploitables. Elle s'inscrit dans une tendance qui considère l'information publique comme un bien commun (tel que défini par Elinor Olstrom) dont la diffusion est d'intérêt public et général. » (Wikipédia, article données ouvertes, consulté le 15/05/2015)

La définition proposée sur Wikipédia reprend les idées d'une donnée numérique en libre accès et réutilisables, cependant elle complète cette définition en précisant que cette donnée doit être structurée, dans sa forme, mais également soumis à une licence définissant les conditions et règles d'utilisation de cette dernière.
Enfin la définition de Wikipédia présente l'intérêt de dépasser le simple fait d'ouvrir les données, mais inscrit le processus dans « un mouvement, une philosophie. Et une pratique de publication »

Des associations de références, françaises et étrangères, dont le travail est reconnu par les institutions ont également proposé des définitions ainsi l'association française LiberTIC qui a notamment assuré la maîtrise d'œuvre et l'animation de la démarche Open Data conjointe Pays de la Loire, Loire Atlantique, Nantes Métropoles et Ville de Nantes propose la définition suivante qui se base sur les dix critères proposés par la fondation Sunlight (association internationale pour la promotion de l'Open Data) :

« Le mouvement « Open Data » (ou « données ouvertes ») est une démarche de publication de données numériques en ligne selon une série de dix critères garantissant leur libre accès et réutilisation par tous. Ainsi, une donnée brute est dite ouverte si elle est publiée de manière: ***Complète, Primaire, Opportune, Accessible, Exploitable, Non-Discriminatoire, Non-Propriétaire, Libre de droits, Permanente, et Gratuite.*** *Cette démarche de publication concerne tous types de détenteurs de données (entreprises, associations, organismes publics) et notamment les collectivités publiques. » (LiberTIC 2010, p1)*

La démarche Open Data vise à mettre à disposition à tous par le biais d'une mise en ligne, des données disponibles au format numérique. Ces données doivent être accessibles et réutilisables par le plus grand nombre. Mais de quelles données s'agit-il ?

2. … utilise la donnée publique …:

Si ces trois définitions nous permettent d'appréhender la démarche d'ouverture des données, seule la dernière propose une réelle définition de quelles données peuvent être éligible à l'ouverture.
Dans le cadre de mon stage je suis amené à travailler sur l'ouverture des données publiques, c'est pourquoi dans cette étude nous ne nous intéresserons pas aux données ouvertes de la sphère privée ou associative, à l'exception des données créées par les organismes exerçant des missions de service public. Nous nous concentrerons sur les données publiques. Mais qu'entend-on par donnée publique ?

Tout d'abords une donnée publique dont la création est financée par l'impôt doit être considérée un bien appartenant à tous et servant l'Intérêt général.

Les données publiques entendues dans le sens de l'Open Data, sont en France définie par la Directive Européenne PSI (Public Sector Information) du 26 juin 2013 ainsi que la loi française n° 78-753 du 17 juillet 1978 dite loi CADA (Commission d'Accès aux Documents Administratifs). La loi CADA définit dans son article premier les données relatives à son application :
*« Sont considérés comme documents administratifs, au sens des chapitres Ier, III et IV du présent titre, quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions et décisions. » (*Loi Cada Article 1)

La mission Etalab précise cette définition des données publiques :
« Les informations ou données produites ou reçues par une autorité administrative dans le cadre de sa mission de service public, publiées par une autorité administrative ou communicables à toute

*personne en faisant la demande. Ces informations doivent être présentées sous un format permettant leur traitement automatisé et leur réutilisation. » (*Etalab 2013, p4).
En outre, *« les informations nominatives, les informations personnelles et les informations protégées par des secrets prévus par la loi (secret de la défense nationale par exemple) sont exclues du champ des données susceptibles d'être rendues publiques, sauf disposition légale ou réglementaire contraire. Les informations statistiques doivent être publiées dans le respect de la loi de 1951, ainsi que de l'article 285 du Traité instituant la communauté européenne, qui définit le secret statistique. »* (Etalab 2013 p 4).

Ces données sont de deux types différents : les données de référence et les données métiers (Feyt, Lardon, 2010). Les données de références se définissent comme faisant « sens communs », elles doivent pouvoir être mobilisées et appréhendées de la même façon quel que soit le territoire ou la thématique de travail, elles doivent présenter un caractère d'homogénéité et de continuité, être le plus complètes possible enfin ces données pour être considérées comme données de références doivent être produites par un organisme de références, légitime qui peut garantir de la qualité de production de cette donnée(Feyt, Lardon,2010). Si l'ouverture le partage, ou l'échange des données de références existait déjà nous l'avons vu précédemment notamment dans le cadre de l'information géographique, l'Open Data favorise l'explosion de la diffusion des données métiers, données qui sont produites dans le cadre des activités des services des institutions et qui n'ont pas pour vocation première d'être échangées distribuées.

On s'aperçoit que la définition proposée par la législation, la mission Etalab et celle proposée par liberTIC se complètent. Dans cette étude nous entendrons donc par données publiques, les données pouvant être ouvertes, produites par une autorité publique ou privée dans le cadre de leur mission de service public, le plus complète possible. Afin de faciliter leur réutilisation et leur traitement par le plus grand nombre ces données doivent être des données brutes mises à disposition dans des formats ouverts. Ces données ne doivent en aucun cas permettre l'identification d'un individu, elles doivent être anonymes ou anonymisées et respecter le secret statistique.

3. …soumise à des licences … :

Le processus même de mise à disposition, de création de la donnée publique, peu la rendre incomplète, non objective, voir non exact. C'est l'une des craintes exposées par les personnes réticentes à l'Open Data. Un autre point qui fait débat est la question du contrôle, de la propriété des données. Pour répondre à cela et pour fixer un cadre à la réutilisation des données, des licences d'utilisation ont été créées.

Il existe de nombreuses licences pour les données ouvertes, certaines beaucoup plus libre que d'autres. Les principales licences utilisées à l'international sont les licences Creative Commons (CC-BY) développées par les outils collaboratifs Wikimédia. Il existe plusieurs niveaux de licences plus ou moins permissifs. La licence Open Government Licence(OGL) utilisée au Royaume-Unis ou la Licence Open Database License (ODbl)

En France les principales licences utilisées sont la Licence Ouverte (LO) créée par la mission Open Data Etalab. Et la Licence ODbl qui est une licence ouverte massivement répandue à l'international. Aux côtés de ces deux licences, une multitude de licences propres, semi-libre, mettant par exemple en place un système de redevance pour une réutilisation commerciale des données sont développés par des entreprises voire des collectivités se revendiquant comme faisant partie du mouvement Open Data.

- La Licence Ouverte d'Etalab :

Créée en 2011, la mission Etalab est chargée de piloter la politique nationale d'ouverture et de partage des données publiques. Outre la mise en place du portail nationale data.gouv.fr. La mission a également développé une licence fixant le cadre juridique lié aux données ouvertes.

Cette licence affirme la libre et gratuite réutilisation des « informations » à des fins commerciale ou non, l'utilisateur peut librement copier, diffuser, adapter, modifier ou extraire une partie de l'information à la seule condition de mentionner la paternité, la provenance de celle -ci. Elle protège également le producteur en précisant qu'il ne peut être tenu pour responsable des inexactitudes présentes au sein de «l'information ». Et précisant qu'il ne peut pas être tenu pour responsable des réutilisations faites de l'information qu'il fournit.

On remarque que le terme « information » est utilisé en lieu et place de « donnée » ceci afin d'être cohérent avec la loi CADA qui parle d'information publique et non de données. Il est à noter que cette licence reste compatible avec les standards des grandes licences internationales type CC BY ou OGL.

- La licence ODbl :

La licence ODbl est une licence de base de données issue du projet opendatacommons de l'Open Knowledge Foundation, adapté pour être applicable en France par l'association Veni Vidi Libri. C'est notamment la licence utilisée par OpenStreetMap.

Comme pour la Licence Libre, la licence ODbl affirme la libre et gratuite réutilisation de la donnée, à des fins commerciales ou non, l'utilisateur est libre de copier diffuser adapter ou modifier ou extraire des parties de la donnée. Cependant à l'instar de la Licence Ouverte l'utilisateur doit mentionner la paternité de la donnée mais également partager aux conditions identiques la nouvelle base de données créée, ou les productions issues de cette base de données. La licence précise aussi que ces versions doivent rester ouvertes qu'elles ne peuvent être soumises à d'autres restrictions que celle compatible avec l'ODbl. La licence ODbl protège également le producteur de donnée vis-à-vis des erreurs ou inexactitude de la base ou des réutilisations qui en sont faites

Cette licence est donc basée sur un partage à l'identique (share alike), on dit qu'elle est contaminante puisque dans le cadre du croisement de données utilisant la Licence Ouverte et la licence ODbl le résultat de ce croisement devra être rendu en licence ODbl et donc rediffusé. Ces licences présentent l'avantage de rendre accessibles à tous les avancées, les innovations réalisées à partir des bases de données d'origine, de plus son caractère contaminant favorise l'ouverture massive des données. Cependant elles sont aussi décriées par certains car jugées moins « libre ». Cette licence favorise l'ouverture des données comme biens communs, en empêchant la mise en place de « clôture » (enclosure) autour de la donnée, les monopoles sur tel ou tel utilisation de la donnée et favorisant les retombées positives pour l'ensemble des utilisateurs et producteurs.

Le choix de la licence sous laquelle ouvrir ses données doit être bien réfléchie, si la licence protège le producteur des éventuels erreurs qui se trouverait dans les bases de données, cela ne doit pas le

dédouaner d'essayer de proposer des données de qualité, à partir desquelles des services ou des réutilisations innovantes pourront apparaître.

4. …et s'inscrit dans une stratégie plus large :

En France il existe peu de recherche sur l'Open Data actuellement, les initiatives sont récentes et les retombés encore mal perçues. Les impacts économiques ont du mal à être définis. Cependant l'ouverture des données publiques politiquement aurait le potentiel de répondre à trois enjeux majeurs :

Tout d'abord un enjeu de transparence de l'action publique et d'information citoyenne, afin d'améliorer les rapports et de rapprocher les citoyens et les institutions. Un autre tenant de cet enjeu est de donner la possibilité au citoyen de devenir acteur de la décision, de co-construire les politiques publiques (Renchi, 2014). De leur donner les moyens de mobiliser et de comprendre les données ouvertes.

Ensuite un enjeu de modernisation de l'action publique. D'une part l'ouverture des données change le rapport des agents, des services des institutions avec leurs données, ceux-ci ne sont plus propriétaire mais producteur de la donnée (OpenDataFrance 2014). Elle doit faire naître une culture de la donnée parmi les agents, les citoyens et les ré utilisateurs et fluidifier les usages et les rapports à la donnée. A terme c'est également une montée en qualité de la donnée publique qui est également visée.

Le dernier enjeu majeur est le développement économique. D'après le rapport Trojette l'Open Data aurait rapporté huit milliards d'Euros de bénéfice au Royaume–Uni dont deux milliards de bénéfices directs. L'ouverture des données doit permettre le développement d'applications, de services innovants, mais aussi le développement et la pérennisation des entreprises déjà existante. Si aujourd'hui le modèle de la « killer App » (une application qui surpasserait toutes les autres dans son domaine) n'est plus d'actualité, il est vrai que de nombreuses applications sont développées grâce à l'Open Data.

La simple démarche d'ouvrir ses données publiques ne peut permettre de répondre à ces enjeux. Elle doit être accompagnée d'une stratégie globale, qui intègre l'amélioration de la qualité des données, l'animation, la sensibilisation tant interne qu'externe aux données et à leurs réutilisations, l'interopérabilité… L'association LiberTIC met en garde contre ce qu'elle appelle le « Data Blues » (voir figure ci-dessous) (LiberTIC 2013) qui est la chute brutale de l'intérêt autour du projet Open Data après l'emballement du début. LiberTIC mais notamment en avant le besoin d'une réelle réflexion pour proposer un projet Open Data de qualité et dépasser les difficultés techniques, juridiques, culturelles et organisationnelles que pose l'ouverture des données publiques.

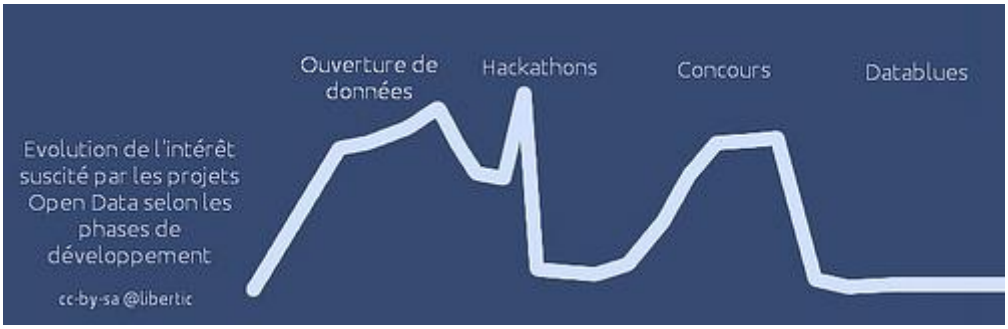


Figure1 : Évolution de l'intérêt de l'Open Data selon les phases de développement

Afin d'ouvrir leurs données, les institutions se dotent de plateformes. Qu'elles soient mutualisées entre acteurs ou relatives à un seul. Ces dernières sont conçues pour permettre l'accès et le téléchargement rapide des données qu'elles reçoivent. Ces plateformes doivent être alimentées régulièrement, les données qui sont dessus doivent être mise à jour, de nouveaux jeux doivent également faire leur apparition. A terme la question de l'actualisation automatique des données se pose, c'est pourquoi de nombreuses plateformes se dotent aujourd'hui d'interface de programmation (API) permettant la mise à jour automatique des données, on a donc un passage des données sous forme de fichier qu'il faut mettre à jour manuellement à des données sous formes de flux. Une fois les données accessibles il faut communiquer pour faire connaître la plateforme et l'animer de façon régulière pour la rendre pérenne.

L'animation vise trois types d'acteurs :

- les ré utilisateurs externes (développeurs datascientist, data journaliste) auxquels il faut proposer régulièrement des animations proposant de travailler sur des sujets concrets, des réunions de travail pour améliorer la plateforme ou des concours pour stimuler la réutilisation des données.
- Les citoyens qu'il faut sensibiliser, auxquels il faut donner les clés pour se saisir de ces données, et impliquer par la suite dans d'autres dispositifs
- Les agents internes qu'il faut également sensibiliser aux intérêts de l'Open Data, les éduqués pour qu'ils changent leur rapport à la donnée.

La frontière entre coups de communication grâce à l'ouverture d'une plateforme et réelle stratégie Open Data est mince, cette animation doit rester régulière. De cette animation émergera une communauté qui constituera un réel acteur avec qui travailler. Au-delà de la portée externe de l'Open Data, c'est un lourd travail de changement de valeur interne qu'il faut mettre en place. A terme l'Open Data ne doit plus être un projet à part entière, mais être une composante de l'ensemble des travaux mener par la collectivité, de 'administration qui met en place cette stratégie.

Conclusion de la première partie :

L'ouverture des données publiques est un sujet récemment apparu au sein des administrations et des collectivités territoriales françaises. L'ouverture des données tire sa source des démarches de partage des données initiées en science, et de la philosophie des acteurs du « libre ». Elle est également liée au développement massif du numérique et de ces usages, dans la société.

En France, l'ouverture des données publiques est encadrée par la loi CADA relative à l'accès aux documents administratifs, mais cette législation est amenée à évoluer. Les données candidates à l'ouverture, sont produites par une autorité publique, ou une entreprise privée dans le cadre de sa mission de service public. Elles doivent être le plus complète possible. Afin de faciliter leur réutilisation et leur traitement par le plus grand nombre ces données doivent être des données brutes mises à disposition dans des formats ouverts. Ces données ne doivent en aucun cas permettre l'identification d'un individu, elles doivent être anonymes ou anonymisées et respecter le secret statistique. Le cadre de leur diffusion est fixé par des licences, ces licences sont plus ou moins permissives, ce qui peut constituer un frein à leur réutilisation.

La démarche d'ouverture doit s'inscrire dans une stratégie plus globale, visant à sensibiliser, éduquer autour de la donnée et animer le dispositif afin de favoriser les réutilisations, mais également à intégrer l'Open Data comme composante de l'ensemble des politiques publiques.

Enfin l'ouverture des données, en rendant accessibles les données métiers, des institutions, ouvre de nouvelles perspectives pour améliorer la connaissance des territoires. Dans la prochaine partie nous allons nous intéresser à comment l'ouverture des données publiques vient enrichir cette connaissance.

II. Open Data et connaissance territoriale :

La donnée constitue la forme initiale permettant de créer l'information, la donnée brute est par nature objective, elle n'est pas encore interprétée. De son interprétation naîtront l'information et la connaissance. Dans le domaine du développement territorial :

« La donnée permet de décrire, d'expliquer et de comprendre le fonctionnement des territoires » (De Sède-Marceau, Moine, Thiam 2009 p9).

Cette connaissance créée est un outil certain d'aide à la décision, qui permet de donner au choix fait un caractère objectif, que l'on ne peut critiquer. Le phénomène d'ouverture des données publiques augmente de façon significative la masse de donnée disponibles et par la même met en avant le concept de la *gouvernance informationnelle des territoires*. (Feyt, Lardon 2010) c'est-à-dire la constitution de l'information comme pilier de la création et mise en place des politiques publiques, la coopération accrue entre les différents acteurs, et disciplines, vectrices de transformations des rapports et formes de gouvernance (Feyt 2011)

A. La donnée comme outil rationnel d'aide à la décision :

1. L'ouverture des données référentielles pour faire sens commun:

Les données référentielles, ou données de référence sujettes à l'ouverture correspondent aux données statistiques et cartographiques qui font sens communs, qui présentent un caractère de continuité, d'homogénéité et qui sont portées par une structure légitime qui permette d'assurer une certaine validité de cette donnée. Cette donnée doit permettre de contribuer à la contextualisation d'informations thématiques, de périmètres ou de métiers particuliers (Feyt, Lardon 2010). Historiquement ces données étaient créées par les grands organismes d'État (IGN, INSEE) ces instituts ont basé depuis leur création leur modèle économique sur la vente des données produites.

Aujourd'hui , de nouveaux référentiels apparaissent que ce soit des initiatives privées issues des géants du web comme Google maps, ou des initiatives collaboratives comme OpenStreetMap, qui met déjà à disposition ses données, l'hégémonie des instituts nationaux disparaît peu à peu et de nouvelles données de référence apparaissent, mieux adaptés aux attentes des citoyens. (Feyt, Lardon 2010). Pour l'État français garder une forme de souveraineté sur ces nouveaux référentiels constitue un vrai enjeu.

La création de la Base Adresse Nationale, projet commun entre l'IGN, la Poste, la mission ETALAB et OSM illustre cet enjeu, en effet jusqu'alors la principale carte permettant une géolocalisation par adresse majoritairement utilisée par les entreprises était la carte proposée par Google maps. L'État ne possédait pas d'équivalent. Un travail conjoint des structures nommées plus haut a permis la constitution de cette base adresse disponible depuis mai 2015. Cette expérience montre également la réticence des grands instituts détenteurs de données, en effet une version libre de cette base de données est diffusée par OSM, une version payante est-elle vendue par l'IGN et la poste.

La décentralisation et l'émergence par le bas de projets de développement remet également en question la pertinence des données de référence classiques. On voit apparaître au gré des projets de nouveaux zonages, de nouveaux territoires, de nouveaux référentiels. Permettre aux acteurs de ces projets de pouvoir se saisir à partir des données de référence existantes, les données dont ils ont besoin pour leur projet sans avoir à payer.

L'ouverture des données référentielles produite par les services publics apparaît aujourd'hui comme nécessaire, en premier lieu pour que ces données restent les données de référence. En effet aujourd'hui les alternatives proposées à ces données payantes sont de plus en plus précises et de bonne qualité. En outre l'avènement d'internet et de la géolocalisation pousse d'une part l'explosion des demandes vis-à-vis de ces données auxquels les institutions ont du mal à répondre, mais favorise également l'émergence de nouveaux référentiels sur lequel l'autorité publique n'a pas le contrôle.

2. L'ouverture des données métiers, et création de connaissance territoriale :

L'autre grande catégorie de données sujette à l'ouverture est la catégorie des données appelées données métiers. Ces données métiers, qui étaient jusqu'à présent peu ou pas accessibles aux acteurs du développement territorial, ouvrent de nouvelles perspectives pour accroître la compréhension des territoires. Si auparavant les données accessibles permettaient la compréhension de phénomène à l'échelle territoriale, l'Open Data rend accessibles les données de gestion, de stocks, de flux, ayant un encrage infra territorial. Ces données développées pour les agents des collectivités territoriales illustrent à une échelle fine les impacts des différentes politiques publiques du territoire. L'accès à ces données ouvre des perspectives nouvelles en matière d'évaluation des politiques publiques et des moyens relatifs à leurs mises en œuvre. Elles ouvrent la voie à de nouveaux modèles de politiques publiques adaptés à une échelle infra territoriale, qui participerait à l'optimisation de ce dernier.

A l'instar de ce qui ce fait en environnement avec la gestion différenciée des espaces, l'ouverture des données publiques pourrait aboutir à une gestion différenciée du territoire. Les politiques publics prendraient en compte les lieux, les temps. On retrouve cette idée dans le concept de smart City (voire figure 2 ci-dessous), à laquelle l'ouverture des données publiques à le potentiel de participer.

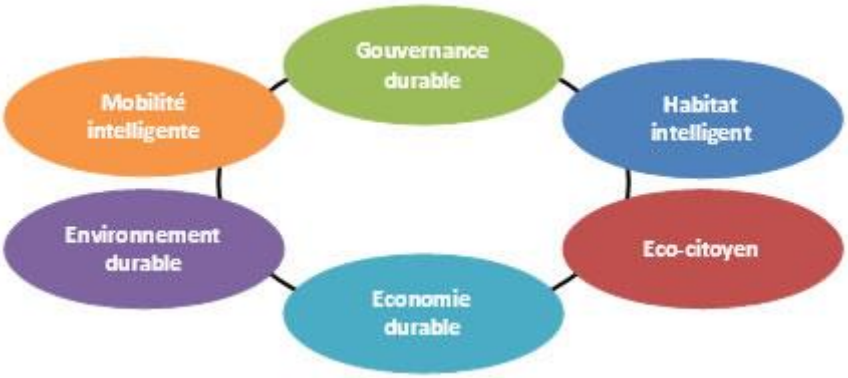


Figure 2 : liste des critères auxquelles doivent répondre les smart Cities (Giffinger 2007)

Ce concept pourrait être étendu on pourrait alors parler de territoires intelligents, qui baseraient leur développement sur une gouvernance participative et l'utilisation efficiente et intégrée des TIC qui

permettrait une gestion des infrastructures, communicantes (entre elles), adaptables, durables et plus efficaces dans le respect des principes du développement durable.

Ces données constituent donc un potentiel nouveau pour faire avancer et évoluer les travaux des chercheurs, des observatoires territoriaux, pour affiner la connaissance territoriale. Mais comment mobiliser ces données pour créer de la connaissance territoriales?

La création de la connaissance :

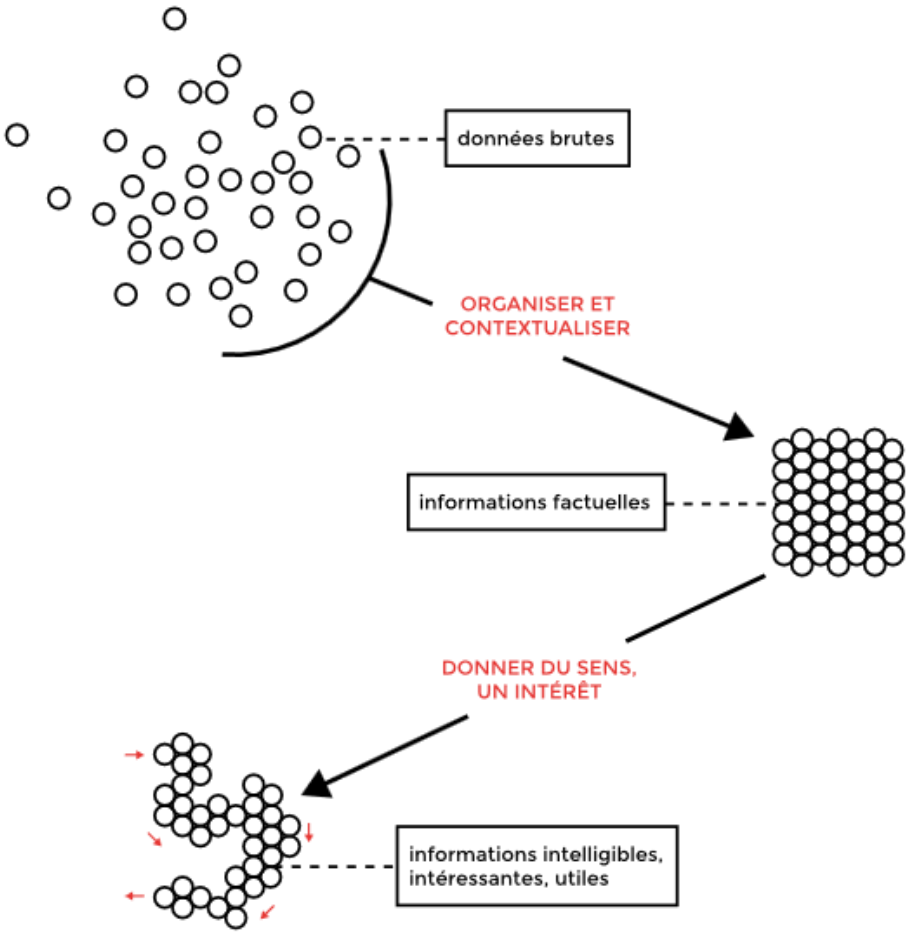


Figure 3 : Processus classique de création de connaissance à partir de données brutes (Czich2015) :

Le processus classique de création de connaissance à partir de données brutes est facilement retraçable, comme le montre la figure 3 ci-dessus en partant de la donnée brute, le résultat d'une mesure, qui est en attente, pas encore exploitée, on applique un premier traitement qui permet de transformer la donnée en information factuelle pour cela elle est organisée, regroupée, avec les données similaires afin de lui faire exprimer une information factuelle, qui exprime quelque chose de réel, à ce stade l'information est encore objective, elle exprime un fait. L'étape qui suit vise à mobiliser cette information factuelle, au service de quelques choses, l'information factuelle devient alors une information intéressante. Le croisement de ces informations mobilisées entre elles et leur analyses, leur mise en relation au regard d'une vision plus sensible, du système de valeurs, des expériences et du savoir de celui qui mobilise l'information crée la connaissance.

La connaissance est donc issue d'une mise en réflexion de l'information.

La création de la connaissance issue des données Open data :

Les données des administrations publiques ne sont pas des entités « déjà là », prédisposées à une ouverture qui tiendrait dans une simple mécanique de libération (Denis, Goëta, 2013). Les données dont disposent les institutions sont généralement issues de logiciel métier, ou ont déjà été traitées pour répondre aux attentes d'un agent, d'un projet d'un service. Le processus d'ouverture de celles-ci passe par trois étapes :

- l'identification du jeu de données au sein des services : c'est l'étape première de toute démarche Open Data, l'institution doit repérer les données dont elle dispose. Bien souvent l'institution n'est pas dotée d'inventaire exhaustif des données qu'elle produit et dont elle dispose, s'engage alors un premier travail de dialogue entre la personne qui est en charge du projet Open Data et les différents services, il s'articule à la fois sur de la communication générale, de la rencontre entre agents et se nourrit de lui-même, c'est-à-dire qu'au fur et à mesure des sollicitations (incessantes) du chargé de mission Open Data, les agents des différents services finissent par eux-mêmes référencer et proposer les jeux de données dont ils disposent. Ce travail permet d'une part d'identifier les données au sein des services mais d'autre part de sensibiliser les agents à l'Open Data.

- l'extraction de la donnée, à partir des logiciels métiers : la plupart des services ne disposent pas de base de données à proprement parler, en effet les données sont « cachées » derrière des logiciels métiers qui fournissent une visualisation et des outils de traitement nécessaires aux usages des agents, mais ces logiciels ne sont pas forcément pensés pour pouvoir accéder aux données qu'ils utilisent. C'est alors un nouveau dialogue qui s'instaure entre le chargé de mission et les informaticiens afin de définir le processus technique (« moulinette » dans le jargon technique) pour extraire ces données et la forme qu'elles doivent avoir en sortie. A chaque nouveau logiciel un nouveau processus. Mais en plus de cela il faut également dialoguer avec les prestataires propriétaires de ces logiciels afin d'obtenir les autorisations de travailler à l'extraction des données et que ce processus ne soit pas considéré comme illégal aux vues des conditions d'utilisation de la donnée.

- La « brutification » c'est-à-dire un traitement permettant sa diffusion et sa réutilisation (Denis, Goëta, 2013) : Nous disposons à ce moment du processus de données extraites des logiciels métiers, il faut maintenant leur donnée un format qui permette leur réutilisation par le plus grand nombre. Le passage d'un format à un autre oblige un travail de nettoyage, d'ajustement afin que les données dans leur nouveau format gardent la même structure, ne soient pas corrompues. Une fois les données disponibles dans un format ouvert, il faut les nettoyer c'est-à-dire retirer les valeurs aberrantes, nulles, supprimer les données liées aux usages spécifiques des agents, mais également toutes formes de mise en page spécifiques, tout ce qui marque ou rappelle les usages précédents de cette donnée. Une dernière étape de ce travail de « brutification » vise à rendre ces données juridiquement ouvrables, c'est-à-dire supprimer les données qui rentre dans le secret statistique, agrégés les données personnelles si cela n'a pas été mis en place lors de l'extraction.

  Les données deviennent alors « brutes » à partir du moment où elles sont déspécifiées, où les données métiers sont transformées pour pouvoir être utilisées à une plus vaste échelle. Ainsi on s'aperçoit que le processus de création de connaissance à partir des données brutes issues de l'Open Data est différent. La donnée subit plus de traitements (voir figure 4 ci-dessous) et l'information qui en découle est donc plus sujette à des choix subjectifs

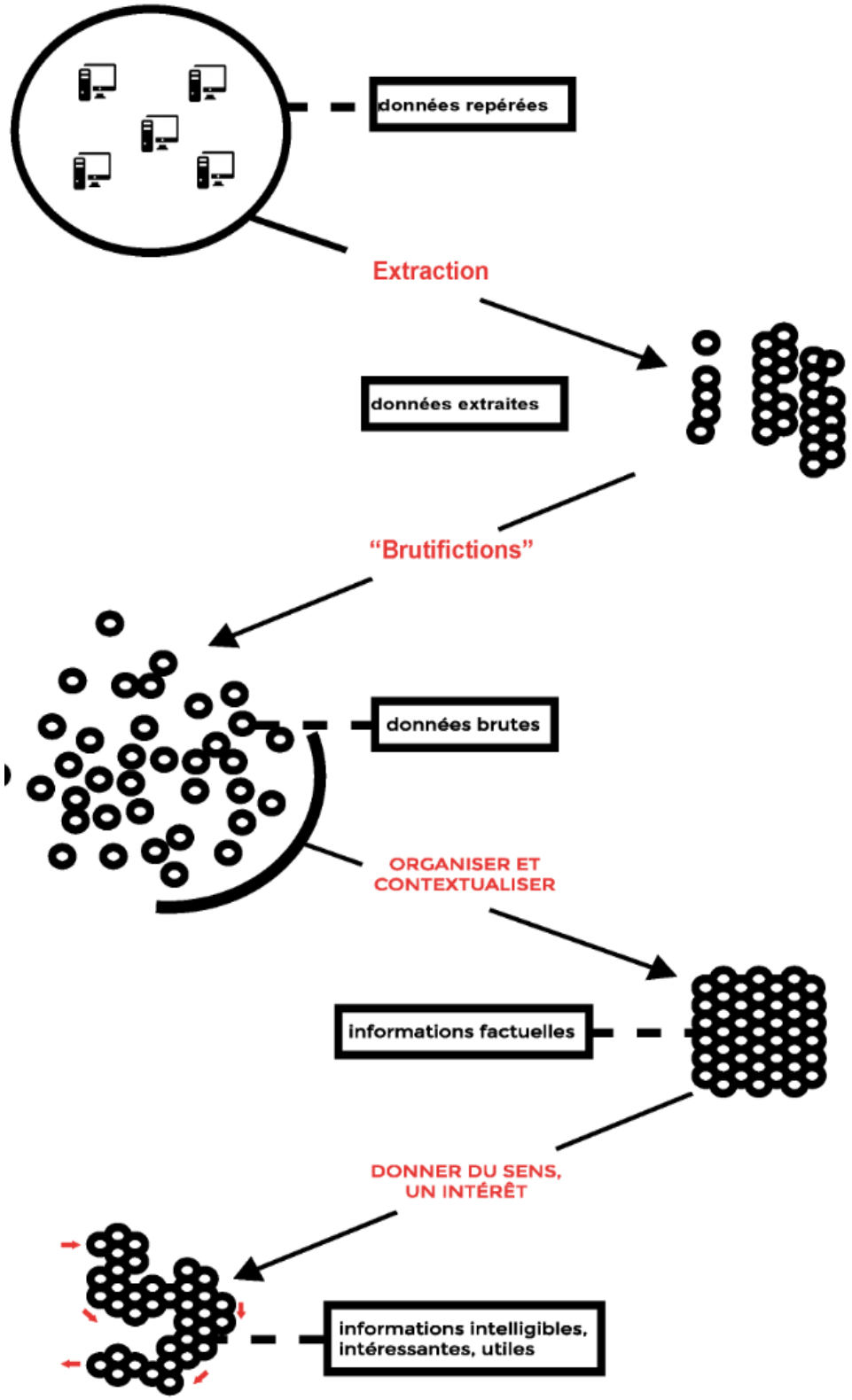


Figure 4 : Processus de création de connaissance à partir des données métiers (Czich, PLATT, 2015):

Ce processus questionne l'objectivité et la validité de la donnée produite, à chaque étape de création de l'information, la donnée subit des traitements issus de choix, qui impliquent une perte potentielle d'information. Il existe de nombreux logiciels différents, toutes les collectivités n'utilisent pas les mêmes et les données qui en sont issues n'ont donc pas la même forme. De plus pour chaque

initiative Open Data créé, les processus d'extraction de la donnée seront différents, aboutissant à la création d'une hétérogénéité de jeux de données relatifs aux mêmes données entre les différents territoires. Il se pose alors la question de l'interopérabilité. Les données brutes de l'Open Data ne s'apparentent donc pas toutes aux données brutes utilisées en sciences ou fabriqué par les SIG.

3. L'interopérabilité des données ouvertes :

L'interopérabilité peut être définie comme *« la capacité que possède un produit ou un système, dont les interfaces sont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre ».* (AFUL 2015)

Dans le cadre des données ouvertes elles consistent à permettre de combiner, d'entremêler différents jeux de données. Dans le domaine de l'information géographique, les questions d'interopérabilité des données sont déjà prises en compte mais continuent d'évoluer. Pour l'ouverture des données publiques, le nombre de dispositifs, de traitement et de processus de mise à disposition des données mène à la création de nombreux jeux de données portant sur les mêmes données mais ayant des formes largement différentes et dispersées dans la multitude de plateforme Open Data mises en place par les collectivités. Rendre visibles, accessibles et interopérables ces jeux de données apparaît comme un enjeu majeur pour leur mobilisation au service de la création de connaissance à une échelle plus étendue et pour favoriser la diffusion des innovations et des traitements qu'un jeu de donnée peut recevoir localement.

A l'échelle européenne, deux directives cadres ont vu le jour dans les années 2000 pour favoriser l'interopérabilité et la mise en place d'une forme de standardisation commune concernant la donnée et ses métadonnées.

- La directive PSI (public Sector information) :

La directive de l'Union Européenne concernant la réutilisation des informations du secteur public du 17 novembre 2003, fixe le cadre général de la diffusion et de la réutilisation des informations du secteur public pour les États membres. Par cette directive, l'Union Européenne met en avant la nécessité de créer des normes de données, pour faciliter la diffusion des documents publics ces données sur la donnée sont appelées métadonnées.

Si cette directive a eu un impact direct sur la normalisation de la métadonnée dans certains pays, par exemple au Royaume-Unis qui a établi un e-Governement Metadata Standard (eGMS), en France elle n'a pas conduit à une création équivalente.

Cette directive a été modifiée en 2013, suite à l'explosion des démarches Open data en Europe, afin de proposer une harmonisation des cadres juridiques des États membres relatifs à ces démarches.

- La directive INSPIRE :

La directive de l'Union Européenne sur l'infrastructure d'information géographique dans la Communauté Européenne du 14 mars 2007 *« élaborée par la Direction générale de l'environnement de la Commission européenne, vise à établir en Europe une infrastructure de données géographiques pour assurer l'interopérabilité entre bases de données et faciliter la diffusion, la disponibilité, l'utilisation et la réutilisation de l'information géographique en Europe. »*(CNIG 2010)

Cette directive élaborée pour répondre en priorité aux enjeux environnementaux (comme lors de l'apparition du terme Open Data en 1995) et la nécessité dans ce domaine de disposer d'une information disponible, de qualité interopérable et accessible à l'échelle de l'Europe afin de pouvoir proposer des réponses globales et appropriées. Si les données environnementales sont visées par cette directive, il en va de même pour les données de référence géographique, au total la directive recense 34 thèmes auxquels elle s'applique.

La directive INSPIRE vise à mettre en place une réelle infrastructure commune pour les données géographiques. Elle insiste sur l'importance de pouvoir combiner facilement et de manière cohérente des informations géographiques provenant de différentes sources à travers l'Europe, et de les partager entre différents utilisateurs et applications (interopérabilité).

Cette directive a été traduite en France, et est aujourd'hui appliquée à tous les échelons territoriaux. Cependant l'ouverture des données dépasse le cadre de cette directive, de nombreuses données ne sont pas soumises à cette directive. C'est pourquoi on a vu émerger une réflexion de la part des initiatives Open data elle-même. Ainsi l'association OpenDataFrance regroupant élus, agents de collectivités et associations citoyennes, travaille sur la question de la standardisation des données, de guide visant à favoriser l'interopérabilité des plateformes et des démarches.

4. Les acteurs, créateur de connaissance territoriale :

Nous l'avons vue l'ouverture des données publiques ouvre de nouvelles perspectives et s'accompagnent de nouveaux questionnements pour l'amélioration de la connaissance territoriale. Nous avons jusqu'ici abordé la création de la connaissance issue de ces données de façon dépersonnalisée, en nous attaquant au processus et aux enjeux. Mais qui peut produire cette connaissance ? Qui va traiter ces données pour les rendre intelligibles et utiles au service de la connaissance ? Il ne s'agit pas ici de dresser une liste exhaustive des acteurs susceptibles de mobiliser cette donnée, mais de nous intéresser à quels acteurs sont susceptibles de la mobiliser pour créer de la connaissance territoriale.

La décentralisation a doté les collectivités locales de nouvelles compétences, faisant de la connaissance un enjeu certains afin de disposer des capacités informationnelles requises pour assumer leurs missions sur des périmètres, des échelles et des thématiques largement inédites dans un besoin de légitimation de leur action(Feyt, 2006) soit un souci d'opérationnalité. Ainsi les collectivités territoriales sont par essence créatrices de connaissance territoriale. Au-delà de leurs missions de gestion, ces dernières produisent un gros travail de mise en connaissance de leur territoire d'intervention, les services et les techniciens participent à fournir aux élus les éléments nécessaires à la prise de décision. Si elles sont les productrices des données que l'on souhaite aujourd'hui rendre accessible au plus grand nombre, l'ouverture des données apparaît également comme un changement profond pour les collectivités. Les démarches Open Data poussent non seulement les services des institutions publics à ouvrir leurs données, mais également à les réutiliser. Cette réutilisation est comparable au processus de « dogfooding » issu du monde de l'entreprise. Le dogfooding est un terme utilisé par les ingénieurs d'une entreprise qui utilisent leur propre produit ou accèdent à leurs données par leurs propres API (interfaces de programmation) publiques. (LiberTIC, 2014). Le dogfooding permet aux agents d'aborder leurs données sous la même perspective que les utilisateurs externes. SI ce processus vise en premier lieu à améliorer la qualité des services proposés par une plateforme. Il permet également d'améliorer les jeux de données proposés, en faisant réfléchir sur leur structure, leur validité. Si les données mises à disposition de

tous sont également utilisées en interne, pour les travaux des techniciens, elles se doivent d'être à jour. Ce processus permet également de repérer de nouveaux jeux de données à ouvrir et favorise la création d'une meilleure information comme l'explique l'association Open Data France : *« Dans l'organisation interne des collectivités locales, au-delà de la production des jeux de données, les effets induits de l'Open Data sont désormais plus visibles sur le volet « réutilisation ». Les échanges de services à services, notamment, se fluidifient. En croisant le jeu de données localisées des permis de construire et le filaire du projet de réseau très haut débit sur un système d'information géographique (SIG), une communauté urbaine a pu identifier en 48 heures les zones d'aménagement urbain stratégiques permettant d'optimiser son plan de déploiement. Dans un cadre de demande classique, la validation puis la mise à disposition des informations, auraient sans doute pris plusieurs semaines. Et l'idée d'un tel croisement n'aurait peut-être jamais été envisagée. Aussi la présence d'une plateforme dédiée à l'ouverture des données accélère-t-elle la sensibilisation des agents et contribue à faire émerger une véritable culture d'usage. Jusque-là, les données étaient peu visibles ou difficiles à exploiter, dans des « silos » protégés. Elles deviennent aujourd'hui visibles, accessibles et sont ainsi mieux exploitées » (OpenDataFrance 2013)*

Ainsi, le fait d'utiliser les données ouvertes pour leurs travaux quotidiens contribue à la monter en qualité des données ouvertes à tous et permet la création d'informations et de connaissances de meilleures qualités.

Afin d'affiner cette connaissance, les collectivités se dotent de plus en plus d'outils d'observation, qu'ils soient thématiques ou territoriaux, ces dispositifs ont pour vocation première de créé de la connaissance (Feyt, Roux 2011). *« Les observatoires territoriaux sont des dispositifs correspondant à une combinatoire de ressources matérielles (outils techniques) et immatérielles (méthodologie, savoir-faire), de pratiques, de moyens organisés et agencés pour une finalité : produire des analyses et des représentations susceptibles de répondre à des problématiques territoriales. Si ces organismes disposent tous, plus ou moins, de capacités de collecte, de stockage de données quantitatives et/ou qualitatives, d'analyses et de production, ils recouvrent dans les faits des réalités plurielles en termes : de moyens techniques et humains, organisationnels et/ou de gouvernance, d'échelles territoriales, et de thématiques ou de problématiques territoriales, qu'elles soient généralistes ou plus spécialisées»*(Roux, 2012). Depuis leur création, ces structures ont toujours mobilisé les données auxquelles elles avaient accès. Si historiquement cette donnée provenait principalement des structures de l'Etat, leur production nous l'avons vu n'est plus son apanage, les collectivités se sont mise à produire, notamment des données géographiques. Enfin par le caractère réticulaire qui existe dans l'organisation de ces structures, les observatoires se nourrissent également largement des données produites par les acteurs qui le compose, mais également des productions d'autres observatoires. L'ouverture des données publiques constitue une source nouvelle d'information pour ces structures. Tout l'enjeu est de savoir comment ces structures réussiront à se saisir de ces données.

Il faut d'ailleurs rappeler que les observatoires, au même titre que l'ensemble des structures exerçant une mission de service public, sont également amener à ouvrir et diffuser plus largement les données qu'ils produisent.

L'un des enjeux de l'ouverture des données publiques est la modernisation de l'action publique et notamment l'implication citoyenne. Si les dispositifs de participations citoyennes ont fleuri ces dernières années (forum hybride, conseil de développement...) l'ouverture des données publiques semble tendre à de nouvelles perspectives. Au motif des transformations souhaitées par les collectivités et les administrations qui se lancent dans l'Open Data, de nouveaux dispositifs voient le jour que ce soit lors d'événements visant à enrichir l'information du territoire (cartoparties), ou en faisant remonter les erreurs et les imprécisions des données ouvertes, le citoyen devient le moteur de la création d'informations de meilleure qualité, c'est ce que l'on appelle le crowdsourcing,

littéralement la « production par la foule », Il s'agit de la mise à contribution des citoyens à la production de données , qui aboutit à la création de valeur, de l'amélioration de l'information territoriale. Cependant, si la coproduction organisée avec les habitants constitue un réel apport à l'information disponible, cela nécessite un lourd travail d'accompagnement et d'animation (OpenDataFrance 2013) et de nouveaux dispositifs restent à inventer.

B. Vers une meilleure mobilisation des citoyens au service de la connaissance territoriale : L'animation de la démarche Open Data.

L'ouverture des données publiques répond à des enjeux de transparence et de modernisation de l'action publique, de participation citoyenne, de développement de l'innovation et de développement économique, l'ensemble de ces enjeux place le citoyen au cœur de l'Open Data. Elle donne aux citoyens de nouveaux outils et moyens de participation. A terme « *Une nouvelle expertise citoyenne pourrait ainsi se développer autour des données. Mieux à même de comprendre l'action publique, les citoyens seraient en capacité de se positionner sur les décisions publiques, voire de participer à la co-élaboration des politiques publiques, pour peu qu'on leur en donne la possibilité »* (décider ensemble, 2012, p30). Si le web 2.0 tend à avoir une approche horizontale : tous les acteurs sont émetteurs et récepteurs, l'information circule de manière circulaire, grâce aux réseaux sociaux, au web participatif, l'information est abondante et facilement accessible. Avec l'Open Data, cette tendance s'inverse : la compréhension et l'interprétation d'un jeu de données nécessitent de multiples connaissances, non seulement dans le domaine ciblé (la démographie, l'économie, etc.), mais aussi en informatique, en statistique… De fait, les plateformes de données publiques ouvertes ont une dimension largement technique, peu de personnes sont réellement capables de mobiliser ces données. De plus les plateformes Open Data ne disposent pas forcément de la visibilité nécessaire pour être mobilisée par ces experts. Enfin cette mobilisation doit être pérennisée, elle ne doit pas être ponctuelle mais aboutir à la création d'une réelle communauté autour de ces dispositifs. L'animation apparait alors comme nécessaire.

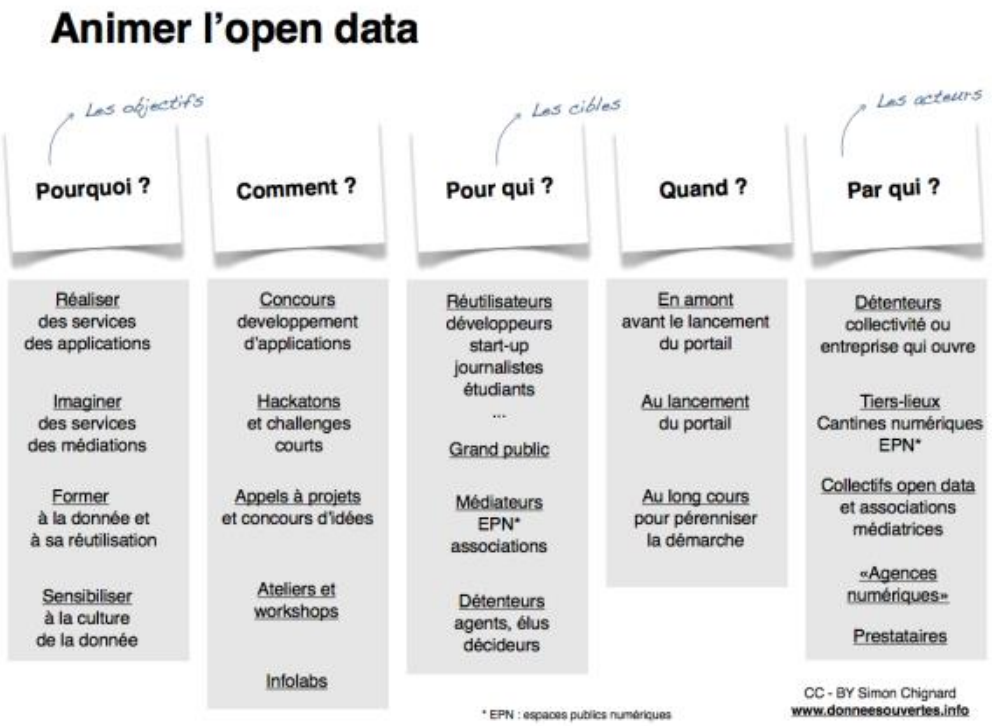


Figure 5 : Grille d'analyse des dispositifs d'animation de l'Open Data Open Data :

Comme nous le montre Simon Chignard par sa grille d'analyse (figure5), l'animation de L'Open Data répond à 5 questions principales, et peut donc revêtir une multitude de formes. Nous allons nous intéresser à l'animation qui place le citoyen au cœur de l'Open Data, que ce soit le grand public ou les réutilisateurs initiés.

1. Mobiliser les citoyens « éclairés » pour créer une communauté :

Comme dans tout projet politique qui se veut phare et innovant, l'attente de résultats visible et quantifiable est très importante. L'Open Data a largement été vendue comme un moteur d'innovation sociale et économique des territoires. A cause de ce besoin de réussite visible et mobilisable, le mouvement Open Data « *érige donc les développeurs en sauveurs de la démarche et sacralise le technicien. Ces « citoyens » capables de donner du sens aux données ont cependant des compétences déterminantes qui les distinguent du reste de la société. »* (LiberTIC, 2012). Ces citoyens que l'on pourrait qualifier d'éclairés sont déjà à même de mobiliser la donnée, de la transformer en information. Cependant il se pose la question de comment les mobiliser ? Si ces personnes sont présentes sur un territoire doté d'une démarche Open Data, il faut d'une part réussir à les repérés, mais également susciter leur intérêt, la mise en place de dispositifs de communications et d'animations apparaît comme nécessaire. Les actions de communication contribuent au rapprochement entre producteurs de données et utilisateurs potentiels. Afin de faire connaître la démarche au plus grand nombre et notamment les réutilisateurs potentiels, il est important de communiquer massivement via les supports traditionnels (médias, réseaux sociaux...) mais également lors de chaque événement organisé par la collectivité (en ouvrant des jeux de données relatifs à cet événement par exemple). Il faut doter la démarche d'une visibilité large. Il s'agit ensuite de faire apparaître l'intérêt auprès des réutilisateurs

L'étude réalisée par l'association « Décider ensemble » témoigne de l'importance des dispositifs d'animation*: « l'animation est un élément clé pour l'émergence d'écosystèmes favorables à la création de valeur économique, la sélection et l'ouverture de jeux de données constituant seulement le premier niveau de la démarche. C'est pourquoi les administrations publiques cherchent à fédérer les compétences autour de communautés ». (déciderensemble, 2012, p42)*.

Par le biais de concours ou de hackatons (Mélange de hack : manipulation ou bricolage d'un système informatique, et marathon), les collectivités stimulent les réutilisateurs, les poussent à mobiliser l'information et à créer des applications, à faire parler les données et créer du service public. « *Ce faisant, il s'agit pour la collectivité locale, de « faire plus avec moins » puisque la foule – en l'occurrence les développeurs militants ou professionnels – est convoquée pour inventer les « services aux publics » de demain hors du service public » (Peugeot, 2011, p11)).*

Mais ces événements ne suffisent pas, s'ils mobilisent de façon ponctuelle et intensive les réutilisateurs, l'émergence d'une réelle communauté utilisant régulièrement la plateforme nécessite d'autres formes d'animations plus régulières. Si les dispositifs d'animation régulière varient dans leurs formes ou dans leur intensités (de l'interaction directe via un forum, à la mise en place de groupe de travail et la création d'instances nouvelles type datalab (voir encadré ci-après), la finalité est la même créer une réelle communauté de réutilisateurs a même de proposer des pistes d'amélioration pour de la démarche, faire remonter les erreurs présente au sein des jeux de données, demander l'ouverture de nouveaux jeux de données.

L'initiative DATALAB en Pays de la Loire :

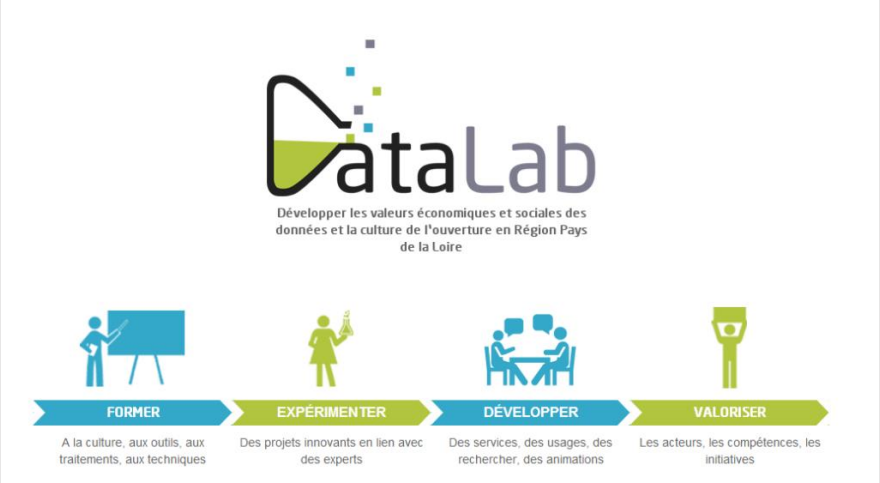


Figure 6 : Dispositif DataLab

Le DataLab est un réseau d'acteurs initié en 2013, il regroupe des collectivités, entreprises, écoles, laboratoires médias, associations… fédérés autour d'un objectif commun **: développer les valeurs économiques et sociales autour des données et de l'ouverture.**
« Ce projet est une proposition de réponse aux limites constatées dans les démarches d'ouverture de données : manque de qualité, diversité et standardisation des données, limite des actions ponctuelles pour soutenir le développement de projets, manque de réutilisation et difficulté d'identification des usages, diversité d'actions sans articulation, difficulté d'appropriation de la culture de la donnée au sein des organisations, freins structurels au développement d'initiatives collaboratives et d'interaction avec l'externe. Dans la première phase de l'open data, il s'agissait d'apprendre à publier des données ouvertes et d'identifier des ré-utilisateurs. Désormais il s'agit également d'apprendre à mieux qualifier les données, diffuser la culture de traitement et de décloisonnement, profiter des échanges externes, améliorer l'accompagnement, identifier des cas concrets de valeur sociale et économique afin de pérenniser les démarches.» (Gallon, 2013)

Claire Gallon présidente de LiberTIC met en avant l'idée de plusieurs phases dans la mise en place de stratégie Open Data, si la première étape vise à repérer et mettre à disposition les données de l'institution, la seconde phase, doit permettre de mobiliser les réutilisateurs pour créer de l'information nouvelle à partir de la donnée. Cette seconde phase de l'Open data induit, le dialogue avec les citoyens que nous avons appelés éclairés, ceux qui savent déjà mobiliser et traiter la donnée, et met la co-construction des données et de leur réutilisation au cœur de la stratégie Open Data, elle doit permettre à une communauté d'émerger. Cette communauté pourra ensuite être mobilisée pour travailler sur des sujets précis, ainsi lors d'événements réguliers appelés Open Data Camp, animés par la mission nationale ETALAB, la communauté de réutilisateurs (comprenant développeurs, techniciens, salariés, étudiants, chercheurs…) est mobilisée durant des ateliers d'une journée sur des thématiques précises. A la différence des hackatons qui sont plus orientés développeurs, ces journées n'ont pas vocation à développer d'applications mais à manipuler la donnée, les combiner, les enrichir. Ce type d'atelier permet de dessiner de nouveaux usages, ou les nouvelles opportunités qu'offre l'accès à ces données.

2. Rendre le citoyen acteur de la connaissance territoriale :

Si la mobilisation des réutilisateurs « éclairés » constitue un enjeu important dans une stratégie d'ouverture des données publiques. Dans un enjeu de transparence et d'implication citoyenne il convient de donner les clés nécessaires à l'ensemble des citoyens pour que ceux-ci puissent se saisir des données ouvertes, les comprendre et les réutiliser. La combinaison ou la confrontation d'une pluralité de savoirs serait bénéfique pour la recherche de la meilleure solution. (Bacqué, Sintomer, 2011). Cet enjeu se rapproche du concept d'*empowerment anglo-saxon soit« le processus par lequel un individu ou un groupe acquiert les moyens de renforcer sa capacité d'action de s'émanciper » (Bacqué, 2005, p30).*

La datavisualisation consiste à transformer une base de données en un élément graphique statique ou interactif afin d'en faciliter son exploration et sa compréhension (Deslis, 2013). Elle permet de rendre intelligible au grand public les données brutes issues des tableaux et bases de données ouvertes, en les transformant en schéma, en cartes, en animation ludique permettant de capter l'intérêt des citoyens. La data-visualisation constitue donc un premier travail de vulgarisation de la donnée. Si ce processus s'avère utile et nécessaire en termes de pédagogie et de transparence de l'information, il ne garantit pas une réelle appropriation. De plus les choix relatifs à chaque data-visualisation correspondent à des interprétations lors de la réutilisation des données et peuvent donner lieu à controverse. En outre, la datavisualisation pose les réutilisateurs de données comme porteur d'un pouvoir d'influence important, les citoyens lambda étant contraints de regarder ce qu'on leur donne à voir et ne pouvant se saisir d'autres de données.

Le simple fait de proposer des visualisations de données grand public ne suffit donc pas, pour que le citoyen se saisisse réellement des données ouvertes il faut l'éduquer, le former. Selon Samuel Goëta et Clément Mabi « *Les enjeux autour de cette formation sont de deux ordres. Le premier consiste à réussir à former l'ensemble de la société aux enjeux de la production et de l'utilisation des données qui « colonisent » le monde social. Cet objectif est atteignable par le développement de dispositifs types « infolabs », espaces hybrides où techniciens et citoyens interagissent pour apprendre à utiliser les données, ou par la mobilisation des espaces publics numériques (EPN) pour réaliser cette « éducation aux données ». Le second consiste à exploiter les données disponibles en open data dans les programmes scolaires pour illustrer par des cas réels des problèmes rencontrés en classe. Ainsi, des données sur l'environnement, la citoyenneté ou encore la géographie sont d'ores et déjà disponibles et pourraient être mobilisées pour fournir des études de cas » (Goëta, Mabi, 2014, p10)* L'utilisation de l'Open Data dans les programmes scolaires constituerait une avancée considérable pour l'apparition de cette « culture de la donnée », cela systématiserait la confrontation des futurs citoyens à ces questions, en outre l'éducation nationale française a pour vocation de favoriser l'émergence d'esprit critique chez ses élèves, dès lors les générations futures disposeraient des clés de lecture et de compréhension de cette donnée. Les dispositifs de types infolabs, sont eux basés sur une démarche volontaire, il faut permettre à ces dispositifs d'avoir de la visibilité, mais également trouver les animateurs nécessaires à ce type de structures.

Faire naitre cette « culture de la donnée » auprès de l'ensemble de la société et des générations futures apparait clairement comme nécessaire, d'une part puisque l'omniprésence et la multiplication des données va continuer de s'accroitre dans nos quotidiens, mais également afin de ne pas accroître la « fracture numérique » (Anderson, 2009) qui ici est entendu plus dans son sens culturel (ceux qui peuvent se saisir de la donnée et ceux qui ne peuvent pas) que générationnel. *« Une prise de conscience par la pratique sensibiliserait au pouvoir des données et apprendrait à ne pas se laisser « éblouir » par les infographies et autres visualisations de données, devenues l'instrument privilégié des communicants » (Goëta, Mabi, 2014, p10)*

L'éducation à la donnée permettrait également de repérer les démarches type « datawashing » : « *similaire à l'éco-blanchiment (greenwashing) le datawashing consiste à publier des jeux de données (présentant souvent peu d'intérêt pour les réutilisateurs) afin de se donner une image de transparence »(Chignard, 2012)*et obligerait les institutions publiques à mieux réfléchir leur stratégie Open Data, d'y intégrer des dispositifs permettant le dialogue avec les citoyens, de prendre en compte leurs demandes en données.

Enfin au-delà de ce travail de sensibilisation, il faut permettre au citoyen de devenir acteur et créateur de la donnée, ce que nous évoquions dans le chapitre précédent, par les démarches types crowdsourcing. Cependant dans la logique de coproduction, et de développement de structures participatives, le citoyen est amené à apporter plus que de simples données, il peut améliorer la compréhension du territoire et de ces processus, raccrocher les réflexions politiques et techniques à une vision plus terre à terre. De fait les productions et les mobilisations de la donnée de la part des citoyens constituent un ajout de valeur différent et complémentaire des productions des techniciens. Ils les poussent à revoir des postures qu'ils pensaient acquises.

Les stratégies Open Data des institutions doivent donc prendre en compte ces enjeux, et permettre le développement de dispositifs qui permettent aux citoyens d'être sensibilisés à la donnée, de mobiliser et croiser ces données, afin de les rendre acteurs du développement de leur territoire, qu'ils participent à l'amélioration de la connaissance territoriale.

3. Les acteurs de la médiation :

Comme nous l'avons vu, la mobilisation citoyenne pourrait permettre de créer de la connaissance et du service public. S'il existe déjà des citoyens « éclairés » qui disposent déjà des compétences pour mobiliser ces données, ces derniers ont besoin d'être stimulée. Une large majorité d'entre eux doit être éduqué afin de leur permettre d'obtenir les clés nécessaires pour qu'ils puissent comprendre et mobiliser la donnée et se saisir des opportunités offertes par l'Open Data. La sensibilisation, l'éducation et la médiation apparaissent comme des démarches nécessaires pour accompagner les citoyens dans ce processus. La démarche d'ouverture des données est donc accompagnée de dispositifs d'animation de natures diverses, nous allons ici présenter les principaux acteurs de cette animation.

Les institutions publiques qui ouvrent leurs données ont bien compris que l'animation des plateformes était nécessaire. Ainsi elles sont assez bien positionnées pour ce qui est de la mobilisation ponctuelle de réutilisateurs, par l'organisation de concours, de hackatons elles favorisent l'utilisation de leur donnée et stimulent la communauté de réutilisateurs. Cette animation est plus tourné vers la création de valeur économique, ou du moins de rendus visibles, concrets, comme des applications mobile que les élus peuvent ensuite utiliser pour illustrer les bénéfices de l'Open Data. Une autre facette de l'animation de la part des collectivités passe par une explicitation claire des données accessibles. Grâce aux métadonnées, ou à des éléments de contextualisations (description de la donnée, mode de collecte…) du jeu de données, ou des systèmes de prévisualisation, la collectivité permet une appropriation plus rapide et plus facile des données. Enfin l'animation de la plateforme auprès du grand public est moins facilement saisissable par les institutions. Le cadre très structuré des administrations ainsi que le manque de moyens humains compétant ne permet pas ou peu la mise en place de projet d'animation d'autres types « *Même si nombre d'administrations publiques lancent aujourd'hui des projets d'accompagnement, cette action, qui est nécessairement une action de terrain, n'entre pas forcément dans leurs missions. » (déciderensemble, 2012, p50).* En outre la question de la neutralité de l'animation, de la sensibilisation se pose, puisque l'administration est à la fois productrice de la donnée, animatrice du dispositif et bénéficiaires des retombés issues de

leur appropriation. Elle fait aussi figure d'autorité ce qui bien souvent peut freiner l'implication de citoyens de plus en plus sceptique.

Ces problèmes se retrouvent dans les dispositifs de concertations, pour y parer, les dispositifs de concertations sont animés par des médiateurs ou « tiers facilitateurs » (Volckrick, 2007), un tiers neutre, indépendants de la maitrise d'ouvrage désigné pour accompagner le dispositif et favoriser le dialogue entre les parties du dispositif. Ils ont un rôle d'interface: *« dans une posture intermédiaire, il transmet la parole du maître d'ouvrage et des parties prenantes, assure la médiation entre les parties prenantes et peut participer à l'animation des débats »* (déciderensemble, 2012, p50). Depuis quelques années, de nombreuses initiatives visant à accompagner le grand public dans la découverte, l'appropriation et la réutilisation des données ont émergées. C'est le cas de l'association LiberTIC, créé en 2009, *« ses actions se développent sous quatre axes: communication, sensibilisations, formations, animations et productions autour des données ouvertes. Parallèlement, l'association participe au mouvement de rapprochement de l'Economie Sociale et Solidaire et des TIC afin de créer des synergies au niveau local et favoriser l'usage des outils libres. »*[1]. Cette association organise de nombreux événements, ateliers de sensibilisation à l'ouverture des données, porteuse de réels projets de concertation, elle est à l'origine du Datalab présenté plus haut. Cette association est souvent mobiliser par les collectivités qui souhaitent se lancer dans une démarche d'Open Data, elle intervient en amont de la démarche pour favoriser l'émergence de bonnes pratiques Cette association est notamment l'un des membres de l'association OpenDataFrance et l'un des co-organisateur de la première semaine européenne de l'Open Data.

Si des associations se sont récemment créées pour répondre spécifiquement à ces enjeux. Des structures plus anciennes se sont également saisies des problématiques liées à l'ouverture des données publiques. La Fondation Internet Nouvelle Génération (FING) créé en 2000, *« elle a pour objet de stimuler, repérer et faire connaître l'innovation dans les services, les applications et les usages de l'Internet de demain. Elle est un projet collectif de veille, de recherche-développement et d'expérimentation. La FING est un projet collectif et ouvert de veille, de recherche-développement et d'expérimentation dont l'objet est de stimuler et faire connaître l'innovation dans les services, les applications et les usages de l'internet de demain. Son ambition est que la France, dans l'Europe, soit un acteur de premier plan de l'innovation dans les usages de l'internet nouvelle génération. Sa conviction est que face aux ruptures qui s'annoncent, une démarche collective d'échange et d'expérimentation est à la fois un facteur d'innovation et un réducteur de risque. »*[2]. Elle s'est donc rapidement saisie des questions d'ouverture des données publiques et à murie une réelle réflexion autour de ces dispositifs. Le fruit de cette réflexion a été le développement du concept d' « InfoLab » (figure 8), sorte de « FABLAB »du numérique et de la donnée. Tout comme les FabLabs, ce concept repose sur des principes généraux tels que l'accessibilité du lieu et l'accompagnement des utilisateurs, l'échange d'expériences et le partage de compétences, l'exploration et le droit à l'erreur. Ce sont des lieux ouverts à tous regroupant le matériel mais également proposant le conseil et le savoir nécessaire à l'appropriation et à la manipulation de la donnée. Véritable lieu d'expérimentation par l'usage, et au-delà de leur mission pédagogique, ces dispositifs favorisent la rencontre entre acteurs et permettent de nouvelles formes d'innovations et la création d'informations différentes.

[1] https://libertic.wordpress.com/libertic/

[2] http://fing.org/

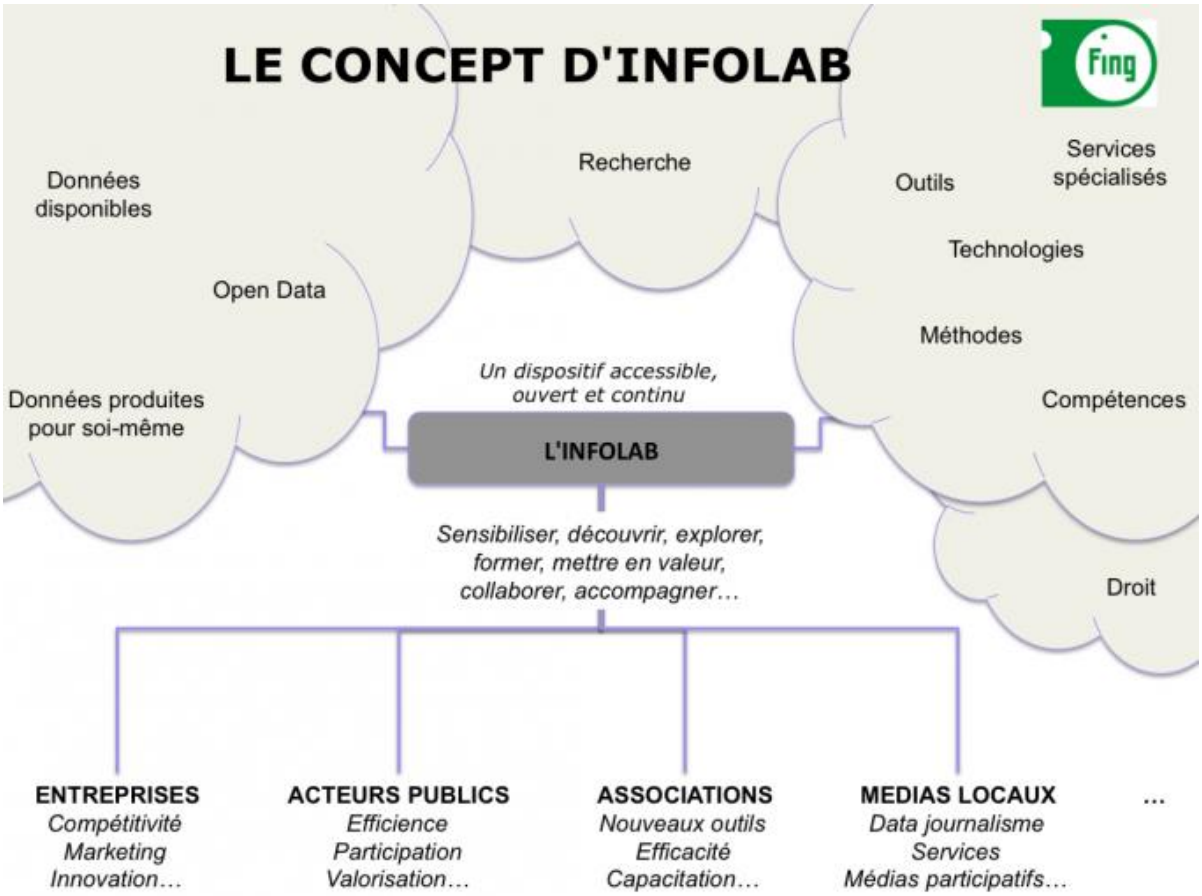


Figure 7 : Le concept d'InfoLab

La mobilisation citoyenne passe par des usages et des codes difficilement appréhendable par les administrations, ainsi la plupart des ateliers ou des rencontres. Elle nécessite de la pédagogie, de la convivialité, du partage. C'est pourquoi la plupart des ateliers, de rencontres, autour des données ouvertes mobilisent des « tiers lieux » des espaces publics informels, neutre, qui favorisent la coopération, entre des acteurs qui ne se rencontreraient pas forcément au quotidien. Ces espaces sont généralement pensés pour mettre en avant le côté convivial, favoriser la discussion, le partage tout en fournissant le matériel et les services nécessaires. (Genoud, Moeckli, 2010). « *Le brassage de compétences, de sensibilités et de points de vue différents mais souvent complémentaires que permettent les tiers-lieux constitue un terreau propice à la naissance d'idées nouvelles au caractère souvent imprévisibles.* » (Genoud, Moeckli, 2010, p5). Ainsi de nombreux tiers lieux sont les supports de l'animation de l'Open Data, les cantines numériques en sont un bon exemple. L'enjeu pour les administrations et les collectivités est alors d'ouvrir le dialogue avec ces structures afin de favoriser la diffusion des innovations, de l'information enrichie.

Conclusion de la deuxième partie :

La transformation des données en connaissance suit un processus de raffinage, dans un schéma que nous avons appelé classique, la donnée brute est mobilisée et contextualisée pour se transformer en information, c'est la mobilisation de cette information, sa confrontation avec l'expérience, le système de valeurs, l'expertise de celui qui la mobilise qui crée la connaissance. Or le processus d'ouverture des données fabrique une nouvelle sorte de données. En fonction des choix effectués aux différentes étapes de ce travail de mise à disposition, les jeux de données créés prennent alors des formes diverses. La nature éparse et morcelée de ces jeux de données est un réel problème pour la connaissance des territoires à une échelle plus large et freine la diffusion des innovations.

Si cette donnée est exploitable et peut venir enrichir les travaux des observatoires territoriaux, des services des collectivités ou des citoyens, il se pose la question de l'interopérabilité entre ces jeux de données.

Nous avons également vu que la démarche Open Data nécessitait une réelle animation afin de pouvoir créer de la connaissance et mobiliser les données. S'il existe déjà des citoyens capables de mobiliser la donnée, ces derniers ont besoin d'être stimulés par l'organisation d'événements ponctuels importants mais également en proposant une animation régulière, autour de la plateforme et de l'ouverture des données. Réussir à créer une réelle communauté permet ensuite de la mobiliser et de faire monter en qualité les données.

Mobiliser les citoyens autour de ces données constitue l'un des enjeux majeurs de l'ouverture des données. Pour cela il convient de les éduquer, leur donnée les clés pour comprendre et être critique vis-à-vis des données. Si les collectivités ont du mal à se positionner sur l'animation auprès des citoyens, des associations qui font aujourd'hui figure de référence, se sont saisis du problème. De plus, des structures nouvelles émergent, ces « tiers lieux » nouveaux espaces de rencontre et d'innovation semblent pouvoir se saisir de ces enjeux.

La création de nouvelles connaissances ne viendra donc pas de faîte, un lourd travail d'animation de la démarche doit être mise en place. Cette animation prend des formes différentes en fonction des publiques et de nouveaux dispositifs restent à être inventés afin de permettre une meilleure implication des acteurs.

III. La Stratégie Open Data de la Région Nord-Pas de Calais :

La réflexion engagée dans les premières parties de ce rapport de fin d'études s'inscrit dans le cadre d'un stage effectué du 15 avril au 30 septembre 2015 à la Région Nord-Pas de Calais. En s'appuyant sur la réalisation de ce stage nous allons illustrer les enjeux mis en relief précédemment et nous tenterons de mettre en avant les limites et point de tension qu'il existe dans la mise en place d'une stratégie Open Data.

Dans un premier temps nous détaillerons le contexte et les attentes relatives à la réalisation de ce stage pour dans un second temps nous intéresser aux réponses proposées et au travail effectuer.

A. Mise en contexte du stage :

1. Présentation de la structure :

La Région Nord-Pas de Calais collectivité territoriale est dotée d'un budget annuel de l'ordre de 2.1 milliard d'Euros. Elle emploie 2265 agents dont 335 agents chargés de missions ou chargés de missions experts travaillant pour les différents services de la Région. La structuration interne de la Région suit la logique suivante, les agents sont regroupés en service dirigé par un chef de service. Ces services sont regroupés au sein de directions dirigées par un directeur. Les Directions sont regroupées pour former des Pôles sous la direction des Directeurs Généraux Adjoints(DGA). Au-dessus de ces DGA il y a la Direction Générale des Services dirigés par le Directeur Général des Services(DGS).

En tant que collectivités elle exerce des compétences obligatoires dans les domaines : du développement économique, de la gestion des transports régionaux de voyageurs(TER), de la formation professionnelle continue et de l'apprentissage, des lycées, de la protection du patrimoine, du développement des ports maritimes et des aérodromes, de la mise en œuvre d'un plan régional pour la qualité de l'air, du classement des réserves naturelles régionales, de l'aménagement numérique, et de la gestion des programmes européens.

Elle est le chef de file pour : l'aménagement et le développement durable du territoire ; la protection de la biodiversité ; le climat, la qualité de l'air et l'énergie ; le développement économique ; le soutien de l'innovation ; l'internationalisation des entreprises ; l'intermodalité et la complémentarité entre les modes de transports ; le soutien à l'enseignement supérieur et à la recherche.

Elle exerce également des compétences facultatives dans les domaines de : la démocratie participative, l'égalité femmes/hommes, du sport, de la culture du tourisme.

Mon stage prend place au sein du service Information Géographique et Analyse Spatiale(IGAS) de la Direction du Développement Durable de la Prospective et de l'Evaluation(D2DPE) rattaché au pôle Appui au Pilotage des Politiques Publiques(APPP).

2. Eléments de contexte :

a) Eléments de contexte généraux :

En France, la législation relative à l'Open Data est en cours de constitution, elle évolue rapidement, si nous l'avons vu les textes sur lesquelles s'appuie l'Open Data sont en place depuis 1978. L'Open Data est au cœur des discussions sénatoriales actuelles de la loi de Nouvelle Organisation Territoriale de la République (NOTRe), de plus le projet de loi numérique devrait également être proposé dans les mois qui viennent. Ces deux projets de loi représentent des changements potentiels pour la stratégie de développement de l'ouverture des données publiques de la Région.

Ce stage prend également place durant un moment historique de la décentralisation : la fusion des Régions Nord-Pas de Calais et Picardie prévue au 1er janvier 2016. Cette fusion qui sera effective le 4 janvier, avec la prise de fonctions du nouvel exécutif (en effet la fusion accompagne les élections régionales prévues en décembre 2015). La fusion des Régions constitue un profond changement dans l'histoire des Régions. Même si les deux futures anciennes Régions disposent de 6ans pour homogénéiser et harmoniser leurs fonctionnements, leurs stratégies et l'ensemble de leurs composantes, cette fusion suscite actuellement de nombreuses interrogations de la part des techniciens. Elle nécessite également leur mobilisation puisque de nombreuses actions et dispositifs devront être opérationnels dès la prise de fonctions du nouvel exécutif.

b) Le développement de l'Open Data en France :

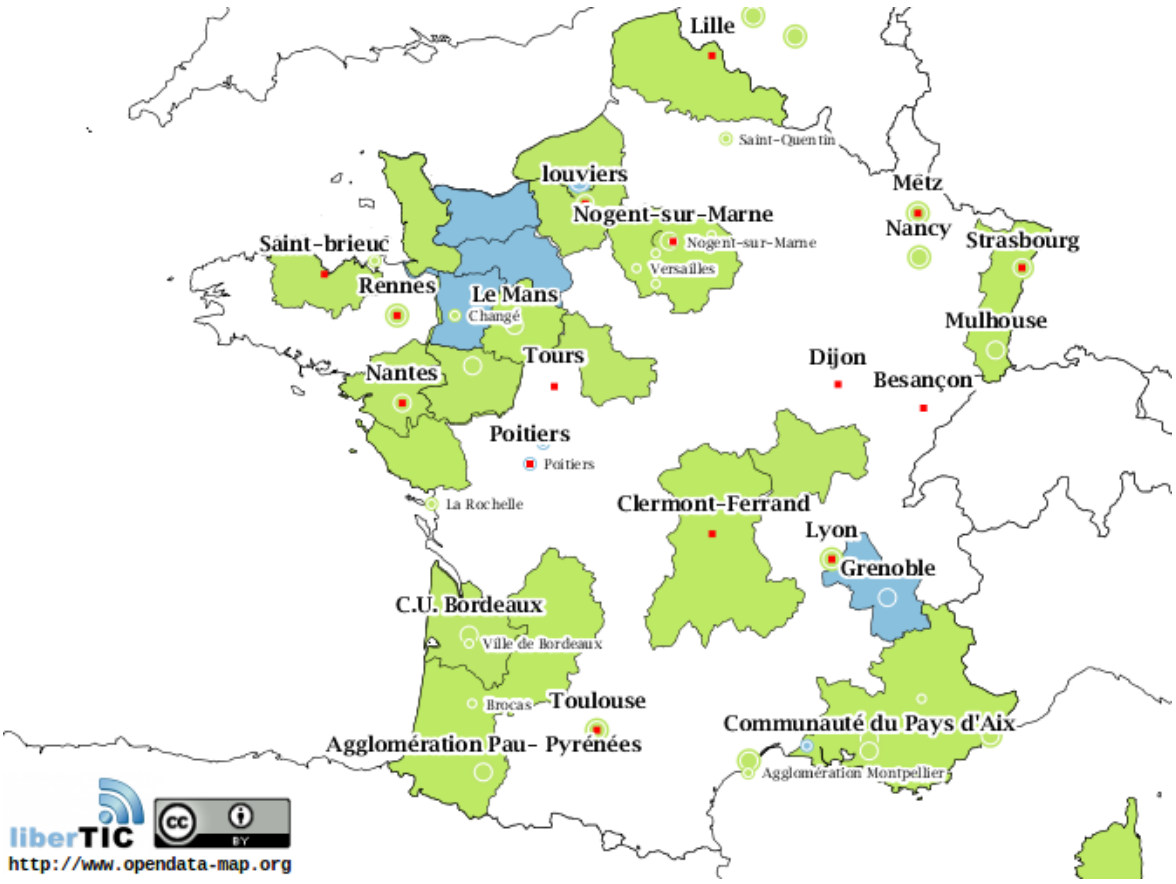
L'Open Data en France d'abord portée par des structures collaboratives associatives (OSM…) est petit à petit devenu un enjeu politique et institutionnel important. C'est en 2010 que les premières démarches d'Open Data institutionnelles voient le jour, la Communauté de Communes Rennes Métropole est la première à se doter d'un dispositif d'Open Data, suivie quelques mois plus tard par la ville de Paris.

A l'échelle nationale, le Gouvernement lance en 2011 la mission ETALAB, chargée de développer et d'animer la plateforme nationale datagouv.fr et plus largement de promouvoir les démarches Open Data. En 2012, dans le cadre de la modernisation de l'action publique, l'Open Data devient un principe gouvernemental, qui prône notamment l'ouverture des données publiques pour une meilleure transparence de l'action publique et l'émergence de services innovants. Dans la lignée de ces démarches un Administrateur Général des données est nommé en 2014.

Aujourd'hui en France, plus de 50 collectivités territoriales de toutes tailles sont engagées dans des démarches d'Open Data (voir figure 8 ci-après). De nombreux mouvements citoyens ont également vu le jour. Une communauté de l'Open Data des collectivités locales est apparue en France, ainsi l'association Open Data France a été créée en 2013.

L'actualité législative récente nous montre que si pour certains secteurs, l'ouverture des données semble être acquise (données transports…), l'ouverture de certains jeux de données suscitent encore un certain nombre de questions et d'inquiétudes (données de santé…).

De plus même si les textes tendent vers l'ouverture des données, les pratiques quant à elles changent de façon moins rapide, il y a encore beaucoup d'interrogation et de méfiance vis-à-vis de l'Open Data



Légende: Déjà ouvert En cours Mouvement citoyen

Figure 8 : Carte des dispositifs Open Data en France

c) Le développement de l'Open Data en Région :

Les premières démarches se rapprochant de l'ouverture de données publiques de la Région ont commencées vers la fin des années 90 avec le développement du Système d'Information Géographique et de l'Analyse de l'Environnement (SIGALE) de la Région Nord-Pas de Calais. Les données géographiques et cartographiques sont rendues disponibles pour d'autres utilisateurs sous une licence spécifique SIG ou un système de convention d'utilisation.

C'est à partir de 2010 que les premières réflexions sur l'Open Data sont lancées en Région, suite au constat de l'inadéquation de la licence utilisée par le SIG régionale avec le « monde libre ». Ainsi en 2011 le choix est fait d'utiliser la licence ODBL (licence libre avec réouverture des données enrichies avec la même licence), puis en 2012 l'Open Data est inscrit au Comité Régional numérique. C'est en février 2013 que la démarche prend réellement son envol avec le vote d'une stratégie d'ouverture des données publiques (Open Data) régionales. Cette stratégie se décline en deux axes principaux :

- Favoriser le développement de l'Open Data parmi les collectivités du territoire, cet axe est porté par la Mission Région Numérique, qui est directement rattachée à la DGS. Concrètement, la Région a décidé pour répondre à ce premier chantier de lancer une démarche d'animation, de sensibilisation des collectivités une assistance à maitrise d'ouvrage (AMO) engagée depuis 2014 pour une durée de 18 mois. Ainsi le cabinet de conseil Beijaflore

anime des ateliers mensuels, auprès des agents des collectivités souhaitant s'engager dans une démarche Open Data. Ces agents peuvent également partager, échanger via la plateforme RECOLTE (REseau CoLlaboratif des acteurs et projets Territoriaux).

- Ouvrir les données du Conseil Régional et en favoriser leurs réutilisations, internes et externes. Cet axe vise à mettre en place une démarche interne d'ouverture des données via la mise en place d'une plateforme permettant la mise à disposition des données de la Région et à son animation.

Ainsi, en 2013, la Région a engagé une démarche expérimentale avec 6 directions pilotes afin de repérer les premiers jeux de données à ouvrir. Cette phase a abouti à la mise en place une première version de plateforme internet dédiée contenant 50 jeux de données régionales. L'ouverture de cette plateforme s'est accompagnée de l'organisation en octobre 2013 d'un séminaire régional rassemblant agents de collectivités et ré utilisateurs privés afin de leur présenter d'une part la démarche initiée par la Région et d'autre part les opportunités offertes par l'Open Data. Ce séminaire a rencontré un franc succès. Cependant, après des débuts prometteurs, la plateforme n'a pas reçu l'animation nécessaire, et l'ouverture des données s'est considérablement ralentie.

Récemment la Région à réaffirmer politiquement et institutionnellement l'importance de l'Open Data en votant fin avril 2015 le Schéma Directeur des Usages et Services numériques (SDUS), document stratégique de développement numérique sur le territoire. L'Open Data dispose dans ce document de sa propre fiche stratégique mais est également la composante de nombreuses autres fiches du document.

3. Commande et réponses proposées :

(Offre de stage : voir annexe 2.)

Le stage proposé, porte sur le second axe de la stratégie Open Data de la Région, à savoir l'amélioration de la démarche Open Data du Conseil Régional. Après des débuts prometteurs, la plateforme n'a pas reçu l'animation nécessaire, et l'ouverture des données s'est considérablement ralentie. Ce stage constitue donc une opportunité afin de redynamiser la démarche.

a) Reformulation :

Il m'est proposé durant ce stage de travailler sur 4 chantiers différents,

Un premier chantier concerne l'enrichissement de la plateforme, c'est-à-dire repérer les jeux de données disponibles en interne et les ouvrir en Open Data. Ce travail d'identification s'effectue aussi bien sur les données déjà en ligne sur la plateforme SIGALE que sur les ressources internes des services. Un travail de tri et de préparation pourra également être effectué afin de proposer des jeux de données réutilisables et compatibles avec les principes de l'Open Data (anonymisation des données, impossibilité de retrouver des individus après croisement de données etc…).

Un second chantier vise à élaborer des pistes d'amélioration de la plateforme. Il m'est demandé d'étudier l'ensemble des opportunités offertes pour proposer un service adapté tant aux producteurs internes qu'aux ré utilisateurs de la donnée.

Le troisième chantier vise à proposer un événement afin d'animer l'Open Data. Libre dans sa forme, il

m'est demandé de définir quel type d'événement, quel type de public visé et de l'organiser. Cet événement doit aussi être pensé de telle façon qu'il puisse être reproduit, reconduit les années suivantes, même si le thème ou les modalités pourront changer.

Enfin le dernier chantier proposé est de travailler à la réutilisation des données libérées en Open Data. Ce chantier constitue un double enjeu. D'une part il servira à légitimer l'ouverture des données, mais d'autre part il aura également une visée pédagogique, « gratifiante » vis-à-vis des agents et de leur travail.

b) Contraintes identifiées :

Les démarches d'Open Data demandent une animation régulière et active, afin de motiver la communauté de producteurs/ré-utilisateurs des données et favoriser l'innovation. Hors la rigueur et la rigidité des procédures de l'Institution peuvent être un frein à cette animation notamment vis-à-vis de la communication autour de l'Open Data ou du temps de réponse aux demandes de la communauté.

Les élections régionales prévues en décembre constituent une contrainte réelle dans l'organisation de l'événement. En effet il est d'une part difficile de lancer de nouveaux projets durant le dernier semestre d'un mandat. Mais d'autre part, en raison de l'interdiction de contribuer au financement de la campagne électorale d'un candidat, les collectivités territoriales doivent respecter certaines règles pour leur communication. Une collectivité n'a pas le droit de financer une action de communication qui ferait campagne pour un candidat. Cette interdiction est très large et concerne également les actions de promotion des actions réalisées qui pourraient avoir un impact électoral. Hors un événement impliquant de la communication externe pourrait être vu comme ayant un impact électoral. Ainsi il faudra être attentif à la forme et au public visé par l'organisation de cet événement mais également à la communication autour de ce dernier.

Le stage s'effectuant en partie sur la période estivale, il se peut que de nombreux interlocuteurs, personnes ressources, agents de la collectivité ne soient pas présents ou disponibles. Ce qui constitue une contrainte dans l'organisation générale du stage. Ainsi les taches ne requérant que peu d'apport extérieur seront favorisées sur la période juillet-aout alors que les taches rassemblant un public plus important seront privilégiées sur la période mai-juin et septembre.

Enfin la multiplicité des démarches et des initiatives Open Data pose le problème de la visibilité des données auprès des utilisateurs et la question de l'interopérabilité des plateformes. Ces deux points devront être largement pris en compte pour le travail sur l'enrichissement et l'amélioration de la plateforme.

c) Atouts identifiés :

Le développement de l'Open Data est porté à la Région par des agents ayant déjà l'expérience et la culture du « libre ». Les membres du service SIG ont par exemple vectorisé et complété le découpage communal de la Région Nord-Pas de Calais pour la communauté OpenStreetMap. Ces agents sont donc des personnes ressources importantes. Les porteurs de la mission Open Data de la Région sont une équipe motivée sur laquelle je pourrai m'appuyer. Il est à noter qu'au sein des autres directions il y a également des personnes motivées par l'Open Data qui m'ont déjà contacté.

L'Open Data s'inscrit parfaitement dans les stratégies de développement de la Région Nord-Pas de Calais. Le Schéma Régional d'Aménagement et de Développement Durable du Territoire (SRADDT), ou encore la démarche de Transformation Ecologique et Sociale de la Région Nord-Pas de Calais(TESR) ainsi que la Troisième Révolution Industrielle(TRI) mettent en avant la participation citoyenne, la transparence et co-construction de l'action publique, le développement d'une région en pointe sur le numérique ou encore l'émergence de services innovants et utiles, autant d'enjeux auxquels le développement de l'Open Data permet de répondre.

Comme le montre la participation au séminaire d'ouverture de la plateforme en octobre 2013, ou les autres démarches d'Open Data de la région (Plateforme Publique de L'Information GEographique (PPIGE)…), la Région peut compter sur un public demandeur, qu'il faudra motiver, auquel il faudra proposer des choses afin de créer une réelle communauté Open Data.

Enfin, L'actuelle plateforme Open Data de la Région est un outil facile d'utilisation, et facilement ré-appropriable, facilitant ainsi l'utilisation des données par le public.

Opportunités :

- S'appuyer sur des projets plus avancés pour proposer des pistes d'amélioration de la plateforme et des pistes d'animation de celle-ci.

- Co construire la plateforme en s'appuyant sur la communauté de ré utilisateurs et leurs attentes.

- Réutiliser les données ouvertes au sein des services de la Région ou directement via des outils de Dataviz implantés sur le site de la Région.

d) Pistes de travail :

Voici les éléments proposés pour répondre à la commande en début de stage :

Animation de la plateforme :

- Dans le but de préparer et d'organiser un événement sur l'Open Data, je propose dans un premier temps de réaliser une recherche documentaire et de réaliser entre 1 et 5 entretiens exploratoires afin de me familiariser avec les différentes techniques d'animations de l'Open Data et de définir le format que devra prendre cet événement, je me donne ensuite un mois pour la formalisation de l'événement puis son organisation.(les invitations pourront être envoyées avant cette période pour des questions organisationnelles et selon la nature de l'événement choisi).
- Au-delà de ce travail événementiel, je propose de rédiger un document synthétique d'une dizaine de pages présentant différentes solutions d'animations régulières de la plateforme.
- Enfin je propose durant l'ensemble de mon stage, de réaliser une veille de l'actualité relative à l'Open Data et de rédiger des notes de synthèses sur ces dernières, pouvant ensuite être intégrées sur la plateforme ou diffusées en interne.

Enrichissement de la plateforme :

- Afin d'enrichir l'offre en donnée de la plateforme, je propose d'identifier l'ensemble des jeux de données disponibles sur la plateforme en ligne SIGALE qui n'ont pas encore été intégrés à la plateforme Open Data.
- Je propose également d'identifier au sein du service SIG les jeux de données disponibles pour être mis sur la plateforme Open Data (un travail de tri et de préparation de la donnée pourra être effectué)
- Je propose de rédiger deux notes synthétiques, l'une à l'intention des prestataires externes producteurs de données pour la Région à intégrer aux marchés publiques précisant la bonne conduite à adopter dans le cadre de création de base de données qui seront par la suite mise en Open Data. L'autre à l'intention des services précisant la marche à suivre et les informations à fournir avec les jeux de données à ouvrir dont ils disposent.
- Enfin je propose sur la durée de mon stage d'intégrer les jeux de données disponibles sur la plateforme Open Data.

Amélioration de la plateforme :

- Afin de proposer des pistes d'amélioration de la plateforme Open Data, je propose d'établir une liste de thèmes (communication, accessibilité…) à améliorer
- Je propose ensuite de réaliser un benchmark autour de différentes initiatives Open Data en France.
- Je propose enfin de rédiger un document comparatif d'aide à la décision d'une quarantaine page synthétisant l'ensemble du travail de benchmark et portant sur l'ensemble des différentes fonctionnalités potentiellement intégrables afin d'améliorer la plateforme.

Réutilisations de la donnée :

- A définir avec les commanditaires

B. Ma mission au sein de la Région Nord-Pas De Calais révélatrice des enjeux et limites des dispositifs Open Data :

1. L'Amélioration de la plateforme :

La plateforme Open Data de la Région Nord-Pas De Calais (figure 9) constitue l'interface permettant la mise à disposition à tous des données de la Région. Elle est développée à partir du logiciel libre CKAN éditeur de plateforme Open Data dédiée. La version actuelle du site est très basique. Les données y sont déposées dans des formats majoritairement ouverts. Un système de classement par grandes thématiques et par catégories permet de structurer l'organisation des jeux de données et de se déplacer plus aisément au sein du site. Elle dispose aussi d'un moteur de recherche interne et de système de prévisualisation pour certaines données.

Figure9 : capture d'écran de la plateforme Open Data de la Région Nord-Pas de Calais

Proposer un site disposant d'une interface claire, facilitant l'emploi et la réutilisation des données mais également leur téléchargement, leur appropriation mais également permettre les retours et la

valorisation des travaux de la communauté constitue un enjeu important pour mobiliser le public, et les réutilisateurs, c'est pourquoi l'un de mes chantiers consiste à proposer des pistes d'amélioration de la plateforme Open Data.

Nous l'avons vu précédemment les démarches Open Data ont fleuri en France et dans le monde depuis 6 ans, certaines démarches sont plus avancées plus abouties que d'autres. C'est pourquoi pour répondre à ce chantier, je propose de réaliser un benchmark comparant les services mis en place par différentes plateformes Open Data, la méthodologie de ce benchmark ainsi qu'une grille d'analyse est en cours de finalisation (mais disponible en annexe 3). Pour mener à bien ce benchmark je souhaite m'appuyer sur la comparaison de 7 plateformes Open Data différentes dont celle de la Région Nord-Pas De Calais par une approche utilisateur, afin d'obtenir plus d'informations je prévois également la réalisation d'entretien semi-directifs avec les personnes en charge du développement de ces plateformes quand cela est possible. Ces entretiens viseront notamment à obtenir des informations sur l'automatisation de la plateforme, son rattachement au système d'information, afin que les données soient mises à jour de façon automatique. Pour cela je vais mettre en place une grille d'entretien au cours de la seconde quinzaine du mois de juin. Les résultats de cette étude seront présentés au mois de septembre dans un document d'aide à la décision d'une quarantaine de page présentant de façon détaillée la démarche et les résultats. Ce document sera accompagné d'une fiche synthétique permettant la comparaison des différents services proposés grâce à une grille d'analyse.

Conjointement à cette démarche tournée vers l'analyse de ce qui se fait ailleurs, je vais également rédiger un document plus court qui recensera les différents bugs, et problèmes rencontrés lors de l'utilisation de la plateforme et son enrichissement.

Limite à la création de connaissance :

Pour le moment la plateforme Open data de la Région Nord-Pas de Calais permet difficilement d'appréhender les réutilisations des jeux de données. Elle ne facilite pas non plus le contact avec les réutilisateurs, L'amélioration de la plateforme constitue un enjeu important dans le processus de création de connaissance, disposer d'un outil proposant un niveau de service adéquat permet de faciliter l'émergence de réutilisations et favorise la création d'une communauté. Cependant cette plateforme constitue une bonne base de travail, un outil presque vierge qui va pouvoir être largement amélioré afin de répondre le mieux aux attentes des dispositifs Open Data.

2. L'enrichissement de la plateforme :

Dans le cadre de sa politique Open Data, la Région souhaite que de plus en plus de jeux de données soient rendus accessibles au public. En outre la Région ne dispose pas d'un réel inventaire exhaustif des données qu'elle produit : certains jeux de données existent sur différents ordinateurs, traitant des mêmes données mais n'ayant pas les mêmes dates de mise à jour. Le travail d'enrichissement de la plateforme, c'est-à-dire la publication de nouveaux jeux de données présente donc un double enjeu, d'une part il répond aux attentes en termes d'ouverture, et de transparence, l'objectif fixé par le comité général de direction étant d'arriver à 100 jeux de données ouverts d'ici à la fin de l'année. D'autre part il s'attache à faire de la plateforme un espace commun qui regrouperait la majorité des jeux de données de l'institution dans leur version la plus récente et la plus complète, les rendant accessibles facilement pour les services de la Région.

Afin de favoriser l'enrichissement de la plateforme mon travail consiste dans un premier temps à repérer les jeux de données disponibles en interne au sein du service IGAS. Le service IGAS est le

service cartographique de la Région, il dispose de nombreuses données issues des différents services de la Région, c'est aussi lui qui gère la plateforme en ligne SIGALE de la Région Nord-Pas de Calais ainsi que le développement de la démarche d'ouverture des données publiques régionales. A mon arrivée dans la structure une majorité des données disponibles sur SIGALE l'était également sur la plateforme Open Data, dans un premier temps je me suis donc attelé à repérer sur la plateforme SIGALE les dernières données qui n'étaient pas encore ouverte (données en jaunes sur la figure 11 ci-après). Au final 14 jeux de données ont été repérés.

	A	B	C	D	E	F	G	H	I	J	K	L
	CODE	TITRE	CATEG ORIE	DESCRIPTION	DATE CREATION	DATE MISE A JOUR	FREQ UENC	PRODU CTEUR	PROPR IETAIRE	LICENCE APPLICABLE	INTERNE_C OMMENTAIR	INTERNE_S E
77	J0076	Espaces à Renaturer fluviaux	8	Espaces à Renaturer fluviaux - Schéma Régionale de Cohérence Ecologique - Trame Verte et Bleue	01/04/2014	01/04/2014	8	1	1	ODBL	SIG réalisé par le service IGAS pour le	D2DPE/IGA
78	J0077	Grands Equipements culturels ou sportifs régionaux et leurs capacités	13	Régionaux ainsi que leur capacité d'accueil (concerts, évènements sportifs ou tourisme d'affaires)	01/10/2014	01/10/2014	9	1	1	ODBL	Equipement s Culturels Régionaux ainsi que	Contrôle de
79	J0078	Fermeture de fillières de CFA	7	rentrés scolaire de Septembre 2014	01/09/2014	01/09/2014	2	1	1	ODBL	de filières de CFA à la	DFI
80	J0079	Ouverture de filières de CFA	7	rentrés scolaire de Septembre 2014	01/09/2014	01/09/2014	2	1	1	ODBL	filières de CFA à la	DFI
81	J0080	Nombre de prêts aux entreprises par Banque et par Territoires	6	Nombre de prêts aux entreprises par Banque et par Territoires	01/01/2014	01/01/2014	9	1	1	ODBL	prêts aux entreprises	DAE
82	J0081	Nombre de repas et Evaluation servis dans chaque restaurant scolaire des E	7	servis dans chaque restaurant scolaire des EPLE	01/03/2013	01/03/2013	9	1	1	ODBL	repas et Evaluation	Contrôle de
83	J0082	Compte Administratif 2013	9	Rapport détaillé en PDF	05/06/2013		2	1	1	ETALAB		DGF/BC
84	J0083	Les véloroutes et voies vertes en Nord-Pas de Calais 2015	19	Localisation géographique des iti	11/05/2015		2	1	1	ODBL		DTR/SIL
85	J0084	Réserves naturelles régionales :	8	Sites naturels classés à l'initiative du Conseil régional. La	01/01/1993	01/04/2014		1	1	ODBL		D2DPE/SIG
86	J0085	territoires de santé :	1	Les territoires de santé sont les territoires de la démocratie	01/11/2010	-		1	1	ODBL		D2DPE/SIG
87	J0086	Zones d'emploi (Zones A) du Nord-Pas de Calais	1;6	Zone géographique où la majeure partie de la population réside et travaille. Le découpage	01/01/1974	01/01/2010		1	1	ODBL		D2DPE/SIG
88	J0087	Schémas d'Aménagement et de Gestion des Eaux (SAGE) du Nord-Pas de Calais	1;8	Le Schéma d'aménagement et de gestion des eaux est un outil	01/01/1992	01/09/2014		1	1	LO		D2DPE/SIG
89	J0088	Plans Climat-Energie Territoriaux (PCET) du Nord-Pas de Calais	1;18	Le Plan Climat Energie Territorial (PCET) est un projet territorial de	01/01/2014	-		1	1	ODBL		D2DPE/SIG
90	J0089	Universités du Nord-Pas de Calais	7	Localisation des universités de la région et caractéristiques	01/01/2013	-		1	1	ODBL		D2DPE/SIG
91	J0090	Politique de la ville - Contrat de projet 2008 - 2013 du Nord-Pas de Calais	1	Communes concernées par la Politique de la Ville 2008 - 2013	01/01/2008	-		1	1	LO		D2DPE/SIG
92	J0091	Bassins versants du Nord-Pas de Calais	8	Périmètre des bassins versants	01/01/1995	-		1	1	ODBL		D2DPE/SIG
93	J0092	Lycées privés du Nord-Pas de Calais	7	Localisation des lycées privés de la région et caractéristiques	01/01/2013	-		1	1	ODBL		D2DPE/SIG
94	J0093	Collèges privés du Nord-Pas de Calais	7	Localisation des collèges privés de la région et caractéristiques	01/01/2013	-		1	1	ODBL		D2DPE/SIG
95	J0094	Régions forestières du Nord-Pas de Calais	8	Présentation des différentes régions forestières *numérisées*	01/01/1986			1	1	ODBL		D2DPE/SIG
96	J0095	Pseudo-cantons du Nord-Pas de Calais	1	territoire. A la différence du	01/01/1994	01/01/2014		1	1	ODBL		D2DPE/SIG
97	J0096	Orientation des pentes en Nord-Pas de Calais	8	rapport au nord	01/01/1996	-		1	1	ODBL	meme jeu	D2DPE/SIG
98	J0097	pentes en Nord-Pas de Calais	8	pourcentage	01/01/1996	-		1	1	ODBL	meme jeu	D2DPE/SIG
99	J0098	Budget Primitif 2014	9	rapport (texte et chiffres) en PDF + tableaux de chiffres associés	20/12/2013		2	1	1	ETALAB		DGF/BC

Figure10: Extrait du tableau de suivi des jeux de données ouverts (vert) et en attente de validation (jaune).

Durant les premiers mois de réalisations du stage j'ai intégré 9 jeux de données ou mises à jour de données validées par le comité de validation. (En vert sur la figure 10)

Dans un second temps, je me suis intéressé aux données disponibles sur le serveur du service. Le service IGAS dispose d'une base de données tenue à jour qui référence l'ensemble de données utilisées pour réaliser de la cartographie au sein du service. Il m'a donc fallu trier cette base de données afin de repérer les jeux potentiellement ouvrables (annexe 4).Au final 23 jeux de données ont été repérés comme disponible sur le serveur et ouvrable et 41 jeux de données ont été repérés comme ayant servi à faire de la cartographie, potentiellement ouvrable mais non disponible sur le serveur.

L'ensemble des jeux de données repérés suit un processus de validation. Si le processus actuel est relativement long. Une proposition de nouveau processus est en cours de proposition (voir encadré ci-après). Ce nouveau processus est présenté pour se conformer aux attentes des directives européennes (PSI et INSPIRE) mais également pour favoriser l'ouverture des données.
La prochaine étape de mon travail consiste à repérer les jeux de données auprès des autres services et des autres directions, ce travail s'annonce long, et nécessite d'aller à la rencontre des autres services, c'est pourquoi je compte profiter des travaux fait pour animer la démarche pour également mobiliser les différents agents afin qu'ils deviennent force de propositions, qu'ils fournissent eux-mêmes leurs jeux de données. Pour faciliter cela je vais rédiger et distribuer une note en interne précisant la bonne conduite à avoir quant à l'ouverture de leurs jeux de données, quelles

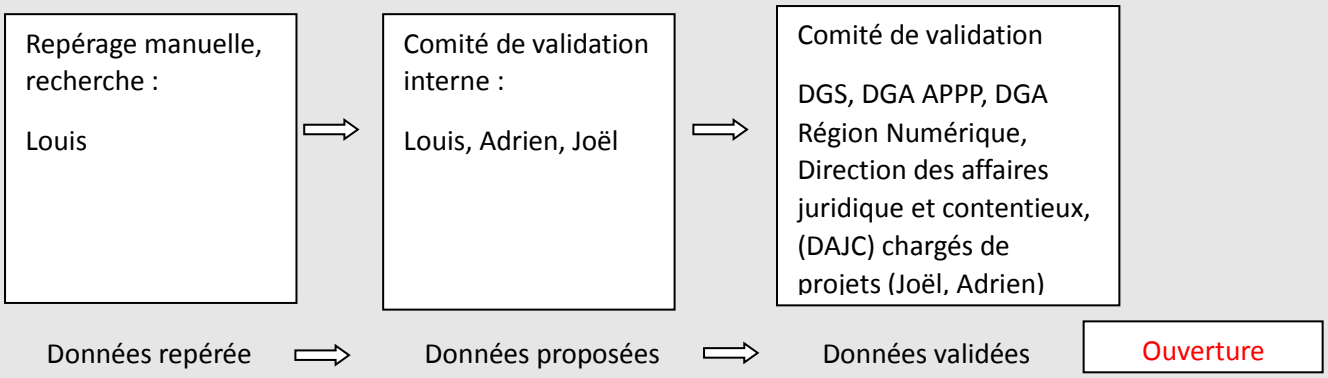
informations fournir avec les données etc… je rédigerai également une note a visé des producteurs précisant la bonne conduite à adopter dans la production de base de données qui seront diffusés en Open Data. Je souhaite attendre que le processus de validation soit validé avant de rédiger ces deux notes.

Parallèlement je continuerai à importer sur la plateforme les jeux de données validés par le comité de validation.

Limite à la création de connaissance :

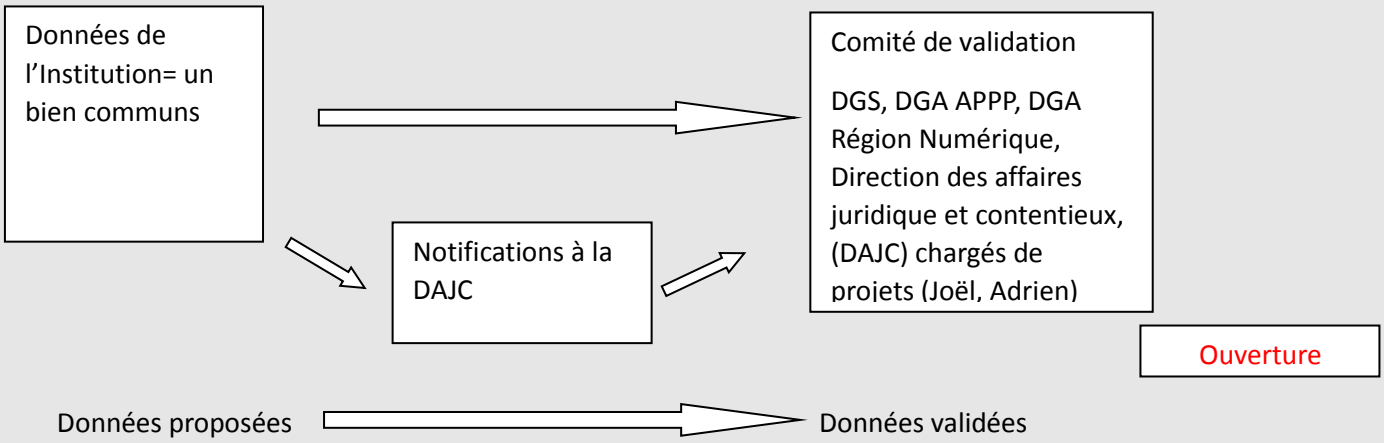
Si proposer un nombre de jeux de données assez conséquent apparait comme un enjeu politique important. La question de la qualité n'en est pas moins primordiale, aujourd'hui en Région le travail de référencement des jeux de données avance, cependant ces données ne disposent pas toujours des métadonnées nécessaires à leur appropriation (même si la mise en conformité des données avec la directive INSPIRE tend à résoudre ce problème). La qualité des données est également un réel problème, en effet le travail de « brutifiction » des données est un processus relativement long actuellement, et qui n'est pas toujours effectué avant la mise à disposition de la donnée (les données sont quand même anonymisées), les données sont mises à disposition « tel quel » ce qui peut être un frein à leur réutilisation.

Processus de validation actuel :



Une fois la donnée repérée, nous en discutons en interne entre chargés de projets et stagiaire, puis les chargés de projets présentent ces données au comité de validation qui vérifie que les données répondent aux exigences juridiques de l'Open Data et ne sont pas des données sensibles, puis valide leur publication. Ce processus de validation actuelle est long, et les nouveaux jeux de données mettent du temps à être validé, en effet le comité de validation ne se réunit pas régulièrement.

Nouveau processus de validation proposé :



Les données produites par les services de l'institution sont considérées comme des Biens Communs, et donc potentiellement ouvrables, sauf avis expressément demandé et justifier auprès de la DAJC. Les données qui arrivent au service IGAS sont de fait mis sur la liste de validation puis présentées au comité de validation qui valide ou non leur publication

3. Animation de la démarche :

L'animation des démarches Open Data est primordiale pour que celles-ci décollent. Si l'animation externe permet de mobiliser les citoyens et les réutilisateurs pour qu'ils se saisissent des données, il est également très importants d'effectuer un lourd travail interne pour faire sauté les a aprioris et mobiliser les agents.

a) Animation externe :

Le travail d'animation externe de la Région passe principalement par l'organisation d'événements, si la phase de lancement de la plateforme a été accompagnée d'un événement rassemblant les foules, ce travail d'animation n'a pas été poursuivit. La rigidité des procédures et les différents niveaux hiérarchiques rendent difficile la mise en place d'animation et de communication régulière auprès des usagers externes.

En Région, les réutilisateurs ne sont pas bien identifiés, il n'y a pas de réelle communauté autour de l'Open Data (Régionale ou à d'autres échelles). Il y a donc ici un gros chantier qui s'annonce. Or le calendrier électoral rend la communication autour des projets de la Région difficile (voir partie contraintes). Ainsi il m'est difficile d'interagir avec l'extérieur, je vais donc sur ce volet réaliser un document proposant différents dispositifs d'animations et les méthodes pour les mettre en marche. Ce document sera basé sur une recherche documentaire ainsi que sur des entretiens avec des personnes-ressources, afin de pouvoir proposer un document méthodologique mobilisable par les chargés de projet après mon stage.

b) Animation interne :

Si l'ouverture des données concerne aujourd'hui 10 directions à la Région, et malgré un portage politique et institutionnel fort, elle ne fait pas encore partie de la culture des agents, les données sont encore considérées comme propriétés des services et des questionnements subsistent. Les principales questions qui nous parviennent sont celles du droit de réutilisations, pourquoi ouvrir les données alors qu'on peut les vendre ? D'autres mettent en avant les erreurs et imprécisions présentes au sein des bases de données pour ne pas les mettre à disposition. De même les réutilisations de la plateforme sont extrêmement rares, aujourd'hui seuls deux services réutilisent des données de la plateforme Open Data. Un lourd travail d'animation et de sensibilisation interne est donc nécessaire.

Pour répondre à ce chantier il m'a été demandé d'organiser un événement, celui-ci prendra place en septembre, sur une demi-journée, j'inviterai une trentaine de personnes sur la base du volontariat à venir, dans un premier temps, assister à une présentation de l'Open Data et de ses enjeux pour dans un second temps travailler collectivement au sein d'ateliers. Ces ateliers ne sont pas encore définis.

Afin de proposer d'animer et de communiquer en interne, j'ai su me saisir des opportunités qui s'offraient à moi ainsi j'ai pu rencontrer lors d'une réunion les référents numériques des différentes directions de la Région. Dans le cadre de son action la mission Région numérique a mis en place des réunions mensuelles qui regroupent des représentants de chaque direction, pour travailler sur les problématiques liées au numérique qu'ils ou leurs collègues sont amenés à côtoyer dans leurs travaux. Nous avons donc participez avec les chargés de projets Open Data, à la réunion du 26mai ou nous avons pu présenter la démarche Open Data et prendre les contacts nécessaires pour pouvoir

repérer plus facilement les jeux de données au sein des services. Cette réunion a fait apparaitre les nombreuses réticences qui existent encore chez les agents.

Dans le cadre de la semaine européenne du développement durable qui se tenait du 30 mai au 5 juin, j'ai été amené à préparer et animer la partie Open Data un stand commun regroupant les démarches collaboratives wikimédias , OpenStreetMap et Open Data régionale, lors d'une rencontre 12/14 portant sur la société du partage (sorte de séminaire ayant lieu le midi, ouvert à tous les agents de la Région, ces séminaires ont lieu dans un cadre moins strict que les réunions classiques et permettent de discuter dialoguer plus facilement, « sandwich à la main » autour de thématiques prédéfini). J'ai notamment préparé un Powerpoint de présentation (annexe 5) et une plaquette de présentation (annexe 6) pour l'occasion. L'animation de ce stand a été particulièrement enrichissante, d'une il m'a permis de rencontrer d'autres agents, de repérer de nouveaux jeux de données à ouvrir. Mais il m'a également fait prendre conscience du besoin de communiquer largement en interne sur l'existence de la plateforme, en effet nombres des agents rencontrés ne connaissaient son existence.

Une autre partie de mon travail d'animation consiste à réaliser une vielle sur l'Open data, cette veille porte principalement sur les évolutions législatives relatives à l'ouverture des données mais également sur les nouveaux dispositifs naissants. Un premier travail de référencement d'articles permet aux chargés de projets de se tenir informer rapidement de l'actualité, lorsqu'ils ont besoin de précision, ce travail de veille donne lieu à la rédaction de notes de synthèses sur certains sujets précis (voir annexe 6), ce travail de veille sera compilé dans un petit dossier rendu en fin de stage, il viendra nourrir la connaissance relative à l'ouverture des données et pourra être diffusé auprès de l'ensemble des agents.

Limites à la création de connaissance :

L'animation joue un rôle primordial pour la mobilisation des données ouvertes, cependant avant ce travail d'animation externe, la rigidité des procédures et les différents niveaux hiérarchiques de l'institution constituent des freins à ce travail de communication et d'animation, il est donc difficile de faire naitre une communauté de réutilisateurs réguliers. En outre un lourd travail d'animation interne doit être réalisé, il faut réussir à mobiliser les agents internes autour du projet. De nombreuses barrières et réticences existent chez les agents, ce travail d'animation interne doit permettre de les faire sauter. La mobilisation des agents doit permettre à termes de récupérer plus facilement les jeux de données, l'aboutissement de ce travail n'étant que ces derniers réutilisent la donnée directement via la plateforme. Pour la démarche de la Région, ce travail en est à ses débuts, beaucoup de frein existe encore chez les agents, le projet ne dispose pas de la visibilité interne nécessaire pour permettre de proposer des jeux de données optimiser pour l'Open Data, ou d'enrichir « naturellement » la plateforme. Ce travail interne constitue un lourd chantier pour le passage de démarche Open Data à stratégie Open data, si la démarche dispose d'un fort portage politique et institutionnel, elle ne dispose pas du même engouement au niveau des agents. Sans cette mobilisation interne, la mise à disposition de données de qualité s'avère être un travail long et difficile et ne favorise pas l'émergence de nouvelles connaissances territoriales.

4. La réutilisation des données :

Le dernier chantier du stage vise à proposer des réutilisations à partir des données mises à disposition sur la plateforme. Ce travail doit servir d'illustration et venir compléter l'animation et la sensibilisation des agents En effet la réutilisations de leurs données permet de mettre en valeur le travail des agents et permet une meilleure mobilisation autour du projet. . Je propose donc de réaliser des cartes et représentations graphiques des données, ces illustrations serviront de support et viendront enrichir l'événement qui sera réalisé au mois de septembre. Cette partie du stage n'est pour le moment pas engagée.

C. Retour d'expérience :

En arrivant en master IDT, je ne savais pas exactement vers quoi me diriger, je souhaitais cependant intégrer un master professionnel qui me permette d'acquérir les connaissances et compétences nécessaires pour me lancer dans la vie professionnelle.

Le choix de la spécialisation géomatique était pour moi nécessaire afin de me doter de compétences techniques « concrètes » et de pouvoir me vendre plus facilement. Après deux ans de master je pense avoir acquis un certain nombre de connaissances en développement territorial et aménagement du territoire, le master m'a également doté d'outils méthodologiques et d'une vision réflexive d'ensemble et capacités analytiques. En outre les différents travaux de groupes réalisés durant la formation m'ont appris à travailler en équipe, à communiquer et à dialoguer avec les autres, autant d'atouts que j'ai été amené à mobiliser durant mon stage.

L'apport des ateliers est un élément réellement important de cette formation, en effet que ce soit en terme de mobilisation de connaissance, de réflexion ou d'organisation du travail il m'a permis d'évoluer rapidement. Avoir l'opportunité de travailler à une commande professionnelle durant la formation est réellement formateur. Il nous permet d'appréhender des facettes ou des thématiques du développement territoriales que nous n'abordons pas en cours, nous responsabilise et nous oblige à mobiliser les outils acquis au service de problématiques qui peuvent nous être étrangères.

Je pense avoir su durant mon stage mobiliser les acquis méthodologiques de ces dernières années, que ce soit dans la mise en place de benchmark, ou l'organisation et l'animation d'un stand, j'ai pu mobiliser mes connaissances, mes compétences et mon savoir-être, sur une problématique que je ne connaissais que peu : l'ouverture des données.

Je souhaitais avant de me lancer dans le monde professionnel pouvoir réaliser un stage en collectivité, appréciant particulièrement l'objet cartographique et l'analyse de données, je souhaitai trouver un stage dans ce champ de l'aménagement. Le stage sur l'ouverture des données publiques de la Région Nord-Pas De Calais m'a donc très vite intéressé. Cela m'a permis d'élargir mes connaissances sur le champ de la géomatique en investissant un domaine que je ne connaissais que peu jusqu'à présent.

Le stage à la Région m'a permis en premier lieu de découvrir et m'intéresser à une problématique que je connaissais peu, depuis ce début de stage j'ai pu acquérir une solide culture quant aux enjeux de l'Open Data et aux éléments à mettre en place pour favoriser son développement. Je suis

désormais à même de dresser le panorama des acteurs et des problématiques relatifs à l'Open Data en France.

En premier lieu ce stage a confirmé ce que je savais déjà c'est-à-dire que la formation que j'avais reçue durant mes années d'études à l'IGA pourrait réellement être mobilisé en milieu professionnel, ce stage m'a permis de confronter les outils méthodologiques et les capacités réflexives développés avec une nouvelle problématique. Mon stage m'a également permis grâce aux agents rencontrés de m'ouvrir au monde du libre et du participatif, domaines auxquels j'étais jusqu'ici éloigné. Au-delà de cela il m'a permis de me retrouver en situation d'animations auprès de professionnels, si je m'en suis bien sorti et que cela constitue une expérience très enrichissante, je me suis rendu compte de la nécessité pour moi de profiter de ce stage pour encore améliorer ce champ de compétence.

Ce stage m'a également permis d'acquérir de solide connaissance dans le fonctionnement de la Région en tant que collectivité territoriale, j'ai été impressionné par la rigidité des procédures et des différents niveaux hiérarchiques qui sont cependant justifiées. Je pense que cette connaissance me servira largement notamment au moment du passage du concours d'attaché territorial que je souhaite passer en 2016.

Conclusion de la troisième partie :

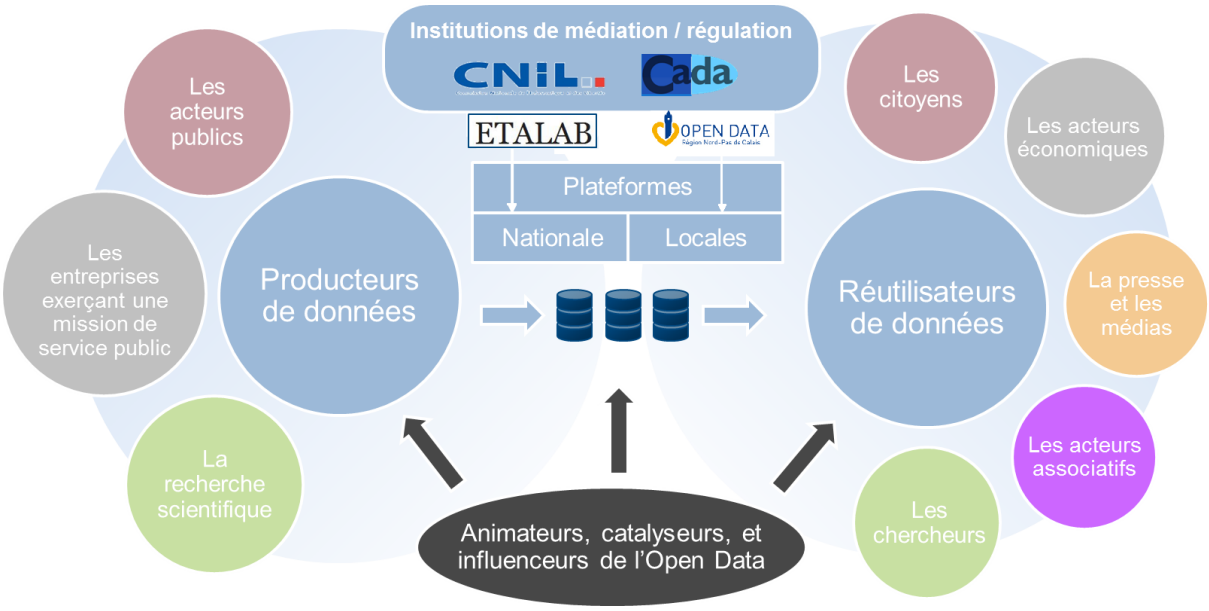
La démarche Open data de la Région Nord-Pas De Calais initié en 2013, s'inscrit dans un phénomène d'ouverture des données publiques de plus en plus répandu à l'échelle nationale. Mon stage s'inscrit dans la poursuite et la redynamisation de cette démarche. Mon travail au sein de la Région m'a permis de me rendre compte des limites et des difficultés rencontrées dans la réalisation d'une démarche Open Data. En outre il m'a permis de me confronter aux différentes composantes de ces démarches à savoir :

- la mise en place d'un outil de qualité : la plateforme,
- le processus de repérage et d'ouverture des données
- la nécessité du travail d'animation.

Cependant si l'Open data ouvre des perspectives nouvelles pour la création de connaissance territoriale, mon stage m'a permis de me rendre compte que la mise en place opérationnelle de cette démarche au sein des collectivités nécessite un lourd travail avant de pouvoir permettre cette création.

Conclusion :

L'ouverture des données publiques est un phénomène récemment apparu en France. Nous avons pu percevoir, à travers cette analyse, qu'elle constituait une nouvelle source d'information. La mobilisation de cette nouvelle donnée constitue un réel enjeu pour l'aide à la décision et la création de connaissances. Nous avons vu que l'ouverture des données tirait son fondement sur les échanges de données déjà présents en science et la philosophie de partage issu du monde du libre. Si l'ouverture est devenue un réel enjeu politique, elle pose de nombreuses interrogations et inquiétude auprès des techniciens notamment en matière de propriété, de droit d'auteur ou encore de qualité de la donnée proposée. L'ouverture des données suit des règles et une législation particulière qui sont amenés à évoluer dans les années à venir. La figure 11 résume parfaitement le schéma de fonctionnement des démarches d'ouvertures en France et rappelle le besoin et la nécessité d'animer l'ouverture.



Source: http://www.digitalcorner-solucom.fr droits de reproduction réservés et strictement limités.

Figure 11: Schéma récapitulatif des démarches Open Data en France

Nous avons également vu que si l'ouverture des données publiques constituait une source nouvelle de donnée, sa mobilisation posait de nombreuses questions. Si la transformation des données en connaissance suit un processus de raffinage, dans un schéma que nous avons appelé classique, la donnée brute est mobilisée et contextualisée pour se transformer en information, c'est la mobilisation de cette information, sa confrontation avec l'expérience, le système de valeurs, l'expertise de celui qui la mobilise qui crée la connaissance. Or le processus d'ouverture des données fabrique une nouvelle sorte de données brutes. La pluralité des démarches et des méthodes de mise à disposition posait alors la question de l'interopérabilité de ces données afin d'étendre analyse et innovations à des échelles plus vastes.

Les acteurs susceptibles de se saisir de cette information étaient de différents types, si en premier lieu les bénéficiaires étaient directement les agents des collectivités et administrations, les observatoires, dispositif fabriquant de la connaissance territoriale seraient à même d'alimenter leurs travaux grâce à ses nouvelles données.

Nous avons également vu que la démarche Open Data nécessitait une réelle animation afin de pouvoir créer de la connaissance et mobiliser les données. Si il existe déjà des citoyens capables de mobiliser la donnée, ces derniers ont besoin d'être stimulés par l'organisation d'événements ponctuels importants mais également en proposant une animation régulière, autour de la plateforme et de l'ouverture des données. Réussir à créer une réelle communauté permet ensuite de la mobiliser et de faire monter en qualité les données.

Mobiliser les citoyens autour de ces données constitue l'un des enjeux majeurs de l'ouverture des données. Pour cela il convient de les éduquer, leur donnée les clés pour comprendre et être critique vis-à-vis des données. Si les collectivités ont du mal à se positionner sur l'animation auprès des citoyens, des associations qui font aujourd'hui figure de référence se sont saisies du problème. De plus, des structures nouvelles émergent, ces « tiers lieux » nouveaux espaces de rencontre et d'innovation semblent pouvoir se saisir de ces enjeux.

La première partie de mon stage m'a quant à elle permit de me rendre compte du lourd travail d'animation à faire en interne avant de pouvoir proposer une plateforme et des données qui puissent être mobilisées au service de la connaissance territoriale. Si l'Open Dat en tant qu'objet politique est un des sujets offrant de nouvelles perspectives, son acceptation et sa mobilisation par les agents et techniciens prendront plus de temps.

Au-delà de cela l'ouverture des données pose également la question de la gouvernance des territoires, si l'un des enjeux de l'ouverture est la transparence, l'information et l'implication citoyenne, de nouveaux modèles de participation restent à élaborer, le citoyen ne doit pas être simple consommateur de données mais devrait également pouvoir dialoguer avec les institutions être impliqué dans la coconstruction des politiques publiques. Là encore, si l'ouverture des données semble pouvoir amener des évolutions dans ce sens, le changement ne se fera que lentement au prix d'un lourd travail.

Annexes:

Annexe 1: Open Government Data Principles

Open Government Data Principles

Government data shall be considered open if it is made public in a way that complies with the principles below:

1. **Complete** : All public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.

2. **Primary** : Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.

3. **Timely** : Data is made available as quickly as necessary to preserve the value of the data.

4. **Accessible** : Data is available to the widest range of users for the widest range of purposes.

5. **Machine processable** : Data is reasonably structured to allow automated processing.

6. **Non-discriminatory** : Data is available to anyone, with no requirement of registration.

7. **Non-proprietary** : Data is available in a format over which no entity has exclusive control.

8. **License-free** : Data is not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Compliance must be reviewable.

a) Definitions

1. **"public"** means:

   The Open Government Data principles do not address what data should be public and open. Privacy, security, and other concerns may legally (and rightly) prevent data sets from being shared with the public. Rather, these principles specify the conditions public data should meet to be considered "open."

2. **"data"** means:

   Electronically stored information or recordings. Examples include documents, databases of contracts, transcripts of hearings, and audio/visual recordings of events.
   While non-electronic information resources, such as physical artifacts, are not subject to the Open Government Data principles, it is always encouraged that such resources be made available electronically to the extent feasible.

3. **"reviewable"** means:

   A contact person must be designated to respond to people trying to use the data.
   A contact person must be designated to respond to complaints about violations of the principles.
   An administrative or judicial court must have the jurisdiction to review whether the agency has applied these principles appropriately

Annexe 2 : Offre de stage

CONTEXTE DU STAGE :
Dans le cadre de sa politique d'ouverture des données publiques (Open Data) votée en février 2013, la Région a mis en place une première version de son site Internet dédié sur lequel 50 jeux de données régionales ont été portés. Elle a également organisé, en octobre 2013, un séminaire régional qui a remporté un franc succès et qui a permis, d'un côté de franchir la première étape-clé du projet régional, et d'un autre d'ouvrir les perspectives sur de futures collaborations avec d'autres collectivités pour aller progressivement vers une infrastructure de données ouvertes.
La première étape-clé franchie, il est maintenant nécessaire de poursuivre l'effort engagé pour aller au bout de la démarche interne : après la mise en place d'un site internet dédié, après l'identification et l'ouverture des 50 premiers jeux de données, après l'organisation d'un événement thématique sur le sujet de l'ouverture des données publiques, il est en effet nécessaire de poursuivre les démarches internes afin :
- d'étendre l'ouverture à un plus grand nombre de directions (6 directions impliquées en première étape) en rencontrant les chefs de service et les directeurs, éventuellement à l'occasion d'un séminaire dédié, et en identifiant des personnes ressources dans les directions ;
- de poursuivre le recensement des jeux de données accessibles et potentiellement ouvrables, et de proposer leur publication au comité de validation des données ouvertes ;
- de veiller à la mise en place progressive des connecteurs entre Systèmes d'Information et back-office du site
web Open Data ;
- de définir les spécifications techniques pour les futures évolutions du site internet Open Data ;
- de concevoir un nouvel événement de communication spécifique à l'ouverture des données publiques et permettant d'associer plus étroitement les ré utilisateurs des données ouvertes.

MISSIONS CONFIÉES AU STAGIAIRE :
Le stage proposé est essentiellement orienté vers les tâches de recensement des jeux de données existants dans les directions, l'identification des fonctionnalités nouvelles à intégrer dans une future version du portail Open Data régional ainsi que vers la conception d'un événement de communication orienté ré utilisateurs.
Les missions confiées au stagiaire sont plus particulièrement les suivantes :
- Participer à l'identification et à la collecte des jeux de données à ouvrir : rencontre des référent(e)s et personnes ressources, présentation de la démarche générale et mise en place du processus de collecte (manuelle) de données, vérification du respect des principes de base de l'ouverture des données pour les jeux de données identifiés et collectés, suivi du tableau de bord en vue de présentation au comité de validation ;
- Participer à la rédaction des spécifications de la version 2 du portail Open Data régional : veille technologique sur le produit open source utilisé par la Région (CKAN), veille sur les sites web similaire, rédaction de propositions d'évolutions fonctionnelles ;
- Participer à l'organisation d'un événement de communication associant les ré-utilisateurs des données régionales : veille sur les types de manifestations existantes, aide au choix d'une formule de communication,
rédaction de propositions et assistance organisationnelle pour la mise en place de cette manifestation, rencontre de ré-utilisateurs

TUTORAT ET RELATIONS FONCTIONNELLES :
Tuteur/trice : Joël TIGNON, chef du Service Information Géographique et Analyse Spatiale et Chef de Projet Open Data
Relations fonctionnelles avec :
-Les différents services de l'institution
- Les autres membres du Service Information Géographique et Analyse Spatiale
- La Direction du Développement Numérique-Systèmes d'Information
- La Direction de la Communication
- Les partenaires extérieures (issues notamment de la sphère des ré utilisateurs de données publiques

PRE REQUIS (formation, expérience, aptitudes personnelles) :
Bac +4/5 dans le domaine de l'Information Géographique, de la géographie, de l'information décisionnelle ou des technologies de l'information
- Appétence pour l'ouverture des données publiques, pour le travail collaboratif sur les données publiques
- Intérêt pour technologies
- Esprit d'ouverture, d'analyse et de synthèse
- Rigueur et organisation dans le travail, autonomie
- Créativité
- Capacité rédactionnelle et sens de la communication
- Capacité à être force de proposition
- Maitrise des outils informatiques bureautiques (Word / Excel / Power Point)

Annexe 3 : méthodologie de Benchmark

Benchmark :

Depuis 2007, dans le monde les différentes instances institutionnelles ont progressivement ouvert des plateformes afin de pouvoir mettre à disposition de tous, leurs données dans un format ouvert. Si les premières démarches ont été initiées par les Etats-Unis, les premières plateformes ont vu le jour en France en 2010. 5 ans plus tard plus de 50 dispositifs de ce genre ont été mis en place en France, à différentes échelles, et sous différentes formes. La Région Nord-Pas de Calais s'est elle aussi lancée dans une démarche d'ouverture de ses données publiques en ouvrant le 8 octobre 2013 une première version de sa plateforme Open Data. Si la plateforme a été enrichie par de nouveaux jeux de données mis à disposition, sa forme quant à elle n'a que peu évolué

Fort de ce constat et afin de proposer des pistes d'amélioration, je propose de réaliser une étude comparative (benchmark) qui permettra de comparer les fonctionnalités et services proposés par un panel de 7 plateformes Open Data différentes de France et de l'étranger.

Méthodologie :

Le Benchmark que nous proposons de réaliser se basera donc sur l'étude des services proposés par 6 plateformes Open Data différentes et de ceux proposés actuellement par la plateforme Open Data de la Région Nord-Pas de Calais. Nous proposons d'élaborer d'organiser notre analyse autour de thèmes et de comparer les différents services proposés par les plateformes pour répondre à ceux-ci. Ainsi pour chaque service étudié nous rédigerons une courte description de celui-ci, établirons les points forts et les points faibles de celui-ci. Chaque description sera accompagnée de captures d'écran et d'illustrations. Enfin à la fin de chaque thème, une grille d'analyse comparative sera remplie permettant de comparer les services entre eux de façon rapide et claire.

Remarque : Les services ne rentrant pas dans cette grille d'analyse mais pouvant présenter un intérêt pour l'amélioration de la plateforme Nord-Pas de Calais disposeront des mêmes fiches descriptives que ceux du benchmark et seront regroupés dans un document annexe.

Les services repérés sur d'autres plateformes et présentant un réel intérêt disposeront également de leur fiche descriptive et seront également compilés dans ce document annexe.

Choix des plateformes à comparer :

En premier lieu, il est important de rappeler que l'ensemble des plateformes observées et comparées dans ce benchmark sont toujours actives et régulièrement mises à jour. Cependant aux vues du nombre d'initiatives Open Data en France et à l'internationale, il m'a fallut sélectionner un panel de plateforme. Le choix du panel de plateforme a été soumis à trois critères principaux[3] :

[3] Les données relatives à ses critères proviennent de www.opendata-map.org/public/opendata-map.csv

Projet collaboratif de cartographie des démarches Open Data en France.

- **La date de création de la plateforme**, en partant du postulat que l'ancienneté de la plateforme a permis aux administrateur de la faire évoluer, de prendre en compte les attentes et les demandes des usagers, ainsi que de fédérer une réelle communauté autour de cette dernière.
- **L'outil web utilisé pour mettre en place cette plateforme**, le choix est fait d'étudier non seulement des plateformes développées sous le même outil que la plateforme Open Data Nord-Pas de Calais (CKAN), mais également d'étudier des plateformes développées avec d'autres outils, afin de regarder d'autres types de services.
- **La question de l'échelle**, en effet si aujourd'hui neuf Régions françaises se sont lancés dans l'Open Data, les premières initiatives datent de 2012. Certains Etats et certaines Communes ou Communautés de Communes se sont saisis de cette démarche bien plus tôt et peuvent proposer des projets plus matures avec des plateformes aux services plus adaptés, auxquels nous devons nous intéresser.

La plateforme Open Data de la Région Nord-Pas de Calais :

https://opendata.nordpasdecalais.fr/

Lancée en octobre 2013, la plateforme Open Data de la Région Nord-Pas de Calais est développée grâce à l'outil libre CKAN, elle contient 80 jeux de données ouverts (au 28/04/2015). La licence utilisée par la Région pour l'ouverture de ses données est la licence ODBL.

Il s'agit de la plateforme pour laquelle nous souhaitons proposer des pistes d'amélioration. Il est donc nécessaire de commencer par évaluer les services proposés par celle-ci.

La plateforme Open Data de la Région Provence-Alpes-Côte d'Azur :

http://opendata.regionpaca.fr/

Lancée en juillet 2012, la plateforme Open Data de la Région Provence-Alpes-Côte d'Azur est la plus ancienne plateforme régionale en France. Elle est développée grâce à l'outil Typo3, et contient 550 jeux de données ouverts dont 132 par la Région (au 28/04/2015). Il s'agit d'une plateforme partenariale, 30 partenaires publics ou exerçants une mission de service public alimentent cette plateforme. L'utilisation de cette plateforme ne soumet pas les partenaires à l'utilisation d'une licence spécifique, ainsi si la majorité des données ouvertes sur la plateforme sont en licence ouverte, certains jeux de données le sont en licence Creatives Commons, d'autres en ODBL et d'autres encore utilisent des licences propres (notamment les jeux de données SNCF et de Nice Métropole)

La plateforme Open Data de la Région Pays de la Loire :

http://data.paysdelaloire.fr/accueil/

Lancée en décembre 2012, la plateforme Open Data de la Région Pays de la Loire est la seconde plateforme régionale à avoir vu le jour en France. Elle est développée grâce à l'outil Typo3, et contient 645 jeux de données ouverts dont 243 de la Région Pays de la Loire (au 28/04/15). Cette plateforme présente la particularité d'être un outil mutualisé entre la Région Pays de la Loire, le département Loire-Atlantique, la Métropole Nantes Métropole et la Commune de Nantes. Chaque collectivité s'appuie sur une plateforme technique commune, avec un catalogue une ergonomie et des services communs, mais dispose d'un site dédié propre, avec une autonomie quant à la ligne éditoriale de ces sites. La licence utilisée pour l'ouverture des données sur cette plateforme est la licence ODBL.

La plateforme Open Data de la Ville de Paris :

http://opendata.paris.fr/page/home/

Lancée en janvier 2011, la plateforme Open Data de la ville de Paris est la seconde plateforme Open Data (toutes échelles confondues) à avoir vu le jour en France. Elle est développée grâce aux outils OPENDATASOFT. Elle contient 129 jeux de données ouverts dont une minorité fournie par des partenaires externes (Centre Nationale de la Cinématographie, Autolib') (au 28/04/2015). Les données publiées sur la plateforme utilisent une licence ODBL, ou une licence adaptée : la licence ODBL Paris.

La plateforme Open Data de la Ville de Montpellier :

http://opendata.montpelliernumerique.fr/

Lancée en mars 2011, la plateforme Open Data de la ville de Montpellier est la troisième plateforme Open Data (toutes échelles confondues à avoir vu le jour en France. Elle est développée grâce aux outils SPIP. Elle contient 212 jeux de données ouverts dont 124 produits par la Ville de Montpellier (au 28/04/2015). Il s'agit d'une plateforme partenariale que 10 partenaires publics mais aussi privés (Open Street Map, et Tela Botanica) alimentent en jeux de données. La majorité des données sont ouvertes sous licence libre, mais il n'y a pas de licence obligatoire sur la plateforme, ainsi certains jeux de données particuliers disposent de licences propres.

La plateforme Open Data gouvernementale du Royaume-Uni :

http://data.gov.uk/

Lancée en janvier 2010, la plateforme Open Data du Royaume-Uni est la première plateforme Open Data à voir le jour en Europe. Elle est développée avec les outils CKAN et DRUPAL. Elle contient 24 536 jeux de données différents (au 28/04/2015), 20440 jeux étant accessibles en Open Data directement sur le site et 4096 accessibles aux personnes autorisées ou soumises à accréditation spécifique car contenant des données non anonymisées ou relative à la sécurité. Cette plateforme est alimentée par 1337 publieurs de données différents, privés ou publics. Les données publiées sont principalement ouvertes sous Open Government Licence (dont la licence ouverte est l'équivalent français), cependant environ 5 000 jeux de données sont soumis à des licences différentes plus ou moins libres. Enfin, il est à noter que le Royaume-Uni est le premier pays au classement mondial de l'Open Data Index de l'Open Knowledge Foundation.

La plateforme Open Data gouvernemental de France :

https://www.data.gouv.fr/

Lancée en décembre 2011, la plateforme Open Data de France est pilotée par la mission Etalab chargé de développer et promouvoir l'ouverture des données publiques en France. Elle est développée à partir des outils CKAN. Elle contient 14 319 jeux de données ouvertes pour 291 contributeurs différents (ayant publié au moins un jeu de données) (au 29/04/2015). Les données de l'Etat et de ses administrations sont publiés en licence ouverte, les données publiées par les autres contributeurs utilisent différents types de licences ouvertes ou non. Il est à noter que la France est aujourd'hui 3ème au classement mondial de l'Open Data Index de l'Open Knowledge Foundation.

Elaboration de la grille d'analyse :

Les études comparatives de plateformes Open Data sont relativement nombreuses, avec chacune une méthodologie, des thèmes et des critères qui leur sont propres. En s'inspirant de 3 études déjà réalisées[4] nous proposons ici une nouvelle grille d'analyse adaptée aux attentes de la commande.

Dans notre grille, les services proposés par les plateformes sont regroupés en 4 thèmes : **données; téléchargement et réutilisation ; communication; animation.** Ces thèmes ont été divisés en sous-thèmes. Afin de permettre la comparaison des services les uns par rapports aux autres, nous positionnerons ces derniers sur une échelle propre à chaque sous thèmes que nous allons détailler ci-dessous :

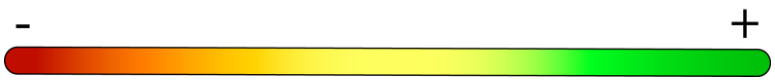
- **Données :**

  - Accessibilité :

    Nous cherchons ici à qualifier l'accessibilité des données qui sont ouvertes par leur format. En regardant le format nous cherchons à juger de la réutilisation potentielle de ces données par le plus grand nombre. Les plateformes proposant une majorité de données ouvertes dans un format ne permettant pas leur réutilisation (PDF…) seront considérées comme moins « accessibles », alors que les plateformes proposant une majorité de données dans un format réutilisable libre seront considérées comme « très accessibles ». Dans le cas où les données sont ouvertes sous différents formats, ce sont les formats permettant le mieux leurs réutilisations qui seront gardés.

    « Quels formats de donnée sont proposés par le service ? »

Majorité de formats non-réutilisable



Majorité de formats réutilisable et ouvert

  - Organisation :

    Ce sous-thème vise à qualifier l'organisation des données au sein de la plateforme. En organisant les données, en les regroupant par thématiques, catégories ou d'autres critères, l'administrateur permet aux utilisateurs non seulement d'accéder rapidement aux données qu'ils cherchent mais également permet un accès plus facile à d'autres données, ainsi les utilisateurs ne recherchant pas une donnée en particulier peuvent aisément voir les données disponibles selon différentes thématiques

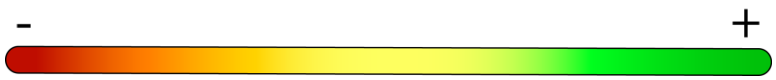
[4] OPENDATASOFT, 2013, Open Data portals Global Review http://fr.slideshare.net/OpenDataSoft/open-data-portals-global-review

LECKO, 2012, L'Open Data au cœur de la nouvelle relation entre collectivités, entreprises et citoyens

OpendataCorisca, HOMER, 2014, guide technique Etat de l'art numerique.corse.fr/file/163637/ et numerique.corse.fr/file/163635/

« Comment les jeux de données sont-ils organisés dans la plateforme ? »

Majorité de formats non-réutilisable



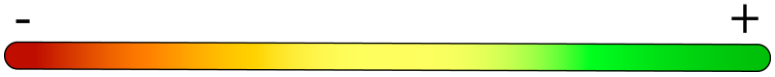
Différentes arborescence structurés permettant de retrouver les données facilement

- Recherche :

Comme le sous-thème précédent, les fonctionnalités de recherches sont là pour améliorer l'expérience de navigation des utilisateurs de la plateforme. Il existe différents outils permettant la recherche : des recherches via des filtres, des moteurs de recherche par mots, des moteurs de recherche avec saisi semi-automatique.

« Que propose les outils de recherche de la plateforme ? »

Pas d'outil de recherche



Outils de recherche multicritères permettant de trier les résultats

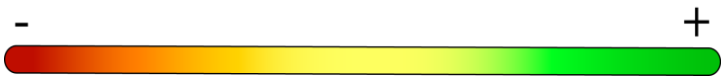
➢ **<u>Téléchargement et réutilisation :</u>**

- Information :

L'information sur la donnée est très importante pour le ré utilisateur, Pouvoir rapidement voir si le jeu de données correspond bien à ce que l'on recherche et contient ce à quoi ont à besoin apparaît comme nécessaire. Ainsi si certains jeux de données peuvent disposer d'une description sommaire présentant globalement le jeu de données en présence, des informations relatives aux champs, aux métadonnées peuvent également être très utiles.

« Quelles informations sont mis à disposition avec les jeux de données ? »

Pas d'information



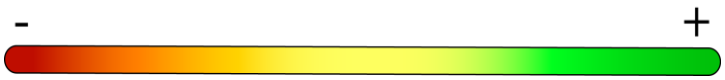
Métadonnées détaillés, listes des champs disponibles et description globale

- Prévisualisation :

Comme le sous-thème précédent, les outils de prévisualisations des données permettent à l'utilisateur de se rendre compte rapidement si le jeu de données qu'il a devant lui est celui qui lui convient. Proposer une ou plusieurs prévisualisations différentes peut améliorer l'expérience de l'utilisateur. Au-delà du simple système de prévisualisation, permettre d'effectuer quelques opérations rapides sur la donnée avant de la prévisualiser constitue l'expérience la meilleure en matière de service de prévisualisation.

« Que propose les outils de prévisualisation de la plateforme ? »

Pas de prévisualisation possible



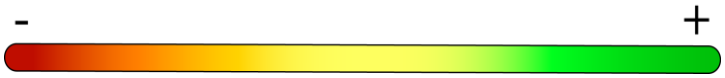
Possibilité de prétraiter la donnée avant de la prévisualiser

- Service aux développeurs :

Faciliter la réutilisation des données aux développeurs constitue un enjeu important dans une stratégie d'ouverture de données. En mettant à leur disposition des API, ces derniers peuvent ainsi faciliter la mise à jour des données utilisés par des applications qu'ils auraient mises en place. De même pouvoir intégrer les remarques et les demandes de ces derniers constitue également un enjeu de taille. Ce sous-thème cherche donc à qualifier la qualité des services proposés pour des ré utilisateurs qualifiés.

« Qu'est ce qui est proposé pour favoriser l'utilisation des données par des développeurs ? »

Pas d'API



Présence d'API et d'outils de consultation communautaire

- Statistiques :
  « Que proposent les outils statistiques de la plateforme ? »

un seul type de statistique	1
plusieurs types de statistiques	2
possibilité de croisés les différentes statistiques des données	3

- **Communication :**

  - Actualités :
    « Comment la communauté de la plateforme est-elle informée ? »

présence d'une rubrique actualité	1
mises à jour régulières des actualités relatives à la plateforme(MAJ de données,concours…)	2
mises à jour régulières relatives à l'actualité Open Data, et de la plateforme)	3

  - Réseaux sociaux :
    « Comment les réseaux sociaux sont-ils utilisés ? »

présence sur un réseau social	1
présence sur plusieurs réseaux sociaux	2
animation régulière des réseaux sociaux de la plateforme	3

  - Réutilisation des données :
    « Comment sont mises en valeurs les réutilisations des données ? »

présence d'une page application	1
mises en valeur des applications par vote de la communauté	2
mises en valeur des applications par vote et commentaire de la communauté	3

- **Animation :**

  - Contact :
    « De quel(s)moyen(s) dispose(nt) la communauté pour contacter le gestionnaire de la plateforme ? »

formulaire de contact	1
contact interactif mais peu animé	2
contact interactif régulièrement animé	3

- Demande, remarques :
  « Comment sont gérées les demandes et remarques ? »

formulaire de contact	1
demandes et remarques publiées sur le site	2
demandes et remarques publiées sur le site avec possibilité d'interaction de la communauté	3

- Evénementiel :
  « Quelle offre en événement accompagne la plateforme Open Data ? »

organisation d'un événement ponctuel	1
organisation régulière d'événements	2
organisation régulières d'événements visant différents publics	3

Annexe 4 : Tableau des jeux de données repérés comme ouvrable au sein du serveur du service IGAS.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Etalab?	Open Data	NOM	CONTENU	Propriétaire	Priorité	politique	Emprise	Commentai	Source(s)	Echelle	Date	Diff.	Red.	Id	Ar	Diffusion		
2	Id Région	A ouvrir	Politique de la ville – Contrat de projet 2007-2013	Communes possédant au moins un quartier inscrit au contrat de projet 2008-	Région					Préfecture de Région / ®IGN - BD Carto®	1/50 000	**2007**	A3	R	npdc_f_vecteur_v401		X	X	X
3	Id Région	A ouvrir	Collèges privés	Localisation des collèges	Région	2	non	NPdC	Voir CG + rectorat + of	Rectorat	1/25 000	**2013**	A3		npdc_f_vecteur_v400		X	X	X
4	Id Région	A ouvrir	Etat des milieux naturels –	Cœurs de nature et cœurs de	Région	1	oui	NPdC	à intégrer	Région	1/25 000	2005	A3		npdc_f_vecteur_v374		X	X	X
5	Id Région	A ouvrir	Schémas d'Aménagem	Dénomination , état	Région	1	non	NPdC	intégré dans bd zonages	Région ; ®IGN - BD Carto®	1/50 000	2006	A3	R	npdc_f_vecteur_v34		X	X	
6	Id Région	A ouvrir	Périmètres de	Périmètre	Région					SNCF / ®IGN -	1/50 000	**2013**	A3	R	npdc_f_vecteur_v338		X	X	X
7	Id Région	A ouvrir	Nature du trait de côte sur le	Nature du trait de côte	Région	3	non	NPdC	A espertiser	Région - ELM	1/5 000	2000	A1		npdc_f_vecteur_v336	X	X	X	
8	Id Région	A ouvrir	Lycées privés	Localisation des lycées	Région	1	non	NPdC	rectorat	Rectorat	1/25 000	**2013**	A3		npdc_f_vecteur_v330		X	X	X
9	Id Région	A ouvrir	Contrats de rivière	Dénomination , état	Région	3	non	NPdC	récupérer data à aeap	AEAP ; ®IGN - BD Carto®	1/50 000	2006	A3	R	npdc_f_vecteur_v33		X	X	
10	Id Région	A ouvrir	Unités cartographiqu		Région / DRAF	3	non	NPdC	à intégrer	DRAF	1/250 000	2005	B1		npdc_f_vecteur_v325		X	X	
11	Id Région	A ouvrir	Bassins versants	Périmètre des grands	Région	3	non	NPdC	vérifier data de l'Agence	®IGN - BD Alti®	1/50 000	1995	A3		npdc_f_vecteur_v28		X	X	X
12	Id Région	A ouvrir	Zones Urbaines Sensibles –	*Communes ayant au moins un*	*Région*					Préfecture de Région / ®IGN - BD	*1/50 000*	*2006*	*A3*	*R*	*npdc_f_vecteu*	*X*			
13	Id Région	A ouvrir	Petites régions	Périmètre	Région					DRAF / ®IGN - BD Carto®	1/50 000	1999	A3	R	npdc_f_vecteur_v191		X	X	X
14	Id Région	A ouvrir	Périmètres de Transport	Périmètre	Région					SNCF / ®IGN - BD Carto®	1/50 000	**2013**	A3	R	npdc_f_vecteur_v190		X	X	X
15	Id Région	A ouvrir	Collèges publics	Localisation des collèges	Région	2	non	NPdC	Voir CG + rectorat + of	Rectorat	1/25 000	**2013**	A3		npdc_f_vecteur_v19		X	X	X
16	Id Région	A ouvrir	Installations internes de	Localisation des	Région	3	non	NPdC	Besoin ponctuel	*DRIRE*	*1/50 000*	*1996*	*A3/A1*		*npdc_f_vecteur_v17*	*X*			
17	Id Région	A ouvrir	Sous zones d'emploi	Périmètre	Région					INSEE / ®IGN - BD Carto®	1/50 000	1983	A3	R	npdc_f_vecteur_v141		X	X	X
18	Id Région	A ouvrir	Zones d'emploi	Périmètre	Région					INSEE / ®IGN - BD Carto®	1/50 000	1983	A3	R	npdc_f_vecteur_v140		X	X	X
19	Id Région	A ouvrir	Sites géologiques	Inventaire des sites	Région	3	non	NPdC	Intégration en l'état	Conservatoir e des sites du	1/25 000	1998	C3		npdc_f_vecteur_v14		X	X	X
20	Id Région	A ouvrir	Circonscriptio ns électorales	Découpage législatif	Région					Préfecture de Région /	1/50 000	**2012**	A3	R	npdc_f_vecteur_v133		X	X	X
21	Id Région	A ouvrir	Végétation Naturelle	Selon la nomenclature	Région	3	non	NPdC	Intégration en l'état	Pr. JM Gehu (1978)	1/250 000	1994	A3		npdc_f_vecteur_v102		X	X	X
22	Id Région	A ouvrir	Régions	Périmètre	Région	3	non	NPdC	Intégration en l'état	IFN / ONF	1/200 000	1994	A3		npdc_f_vecteur_v101		X	X	X
23	Id Région	A ouvrir	Installations de collecte et	Localisation des	Région	3	non	NPdC	Besoin ponctuel	*ARE / DDE59 /*	*1/25 000*	*2006*	*A3*		*npdc_f_vecteur_v1*	*X*			
24	Id Région	A ouvrir	Universités	Localisation des	Région	1	oui	NPdC	of etalab	Région - D2DPE	1/25 000	**2013**	A3		5d00d435-d5f3-4fa2-		X	X	X

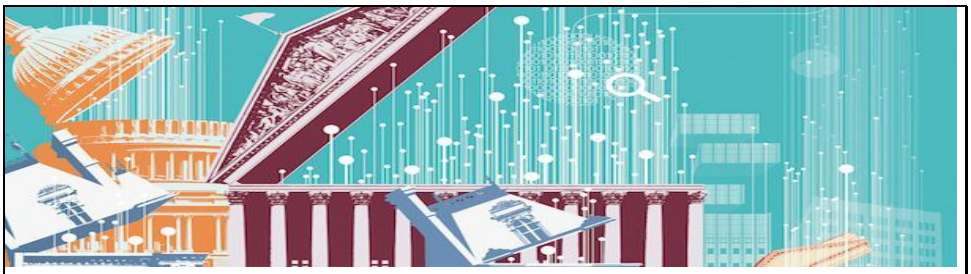
Données disponibles sur le serveur et potentiellement ouvrables

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Etalab?	Open Data	NOM	CONTENU	Propriétaire	Priorité	politique	Emprise	Commentai	Source(s)	Echelle	Date	Diff.	Red.	Id	Ar	Diffusion		
25	Non	A ouvrir	Espaces		Région	1	oui	*NPdC*	à constituer										
26	Non	A ouvrir	espaces info-		Région	2	oui	NPdC	à constituer										
27	Non	A ouvrir	Antennes		Région	3	oui	NPdC	à constituer										
28	Non	A ouvrir	learning			2	oui	NPdC	à constituer										
29	Non	A ouvrir	Equipements	équipements		3	non	NPdC	à constituer										
30	Non	A ouvrir	Chemins de		Région	3	non	NPdC	A intégrer	Région -	1/50 000	2006	A2						
31	Non	A ouvrir	Boucles		Région	3	non	NPdC	A intégrer	Région -	1/50 000	2006	A2	R					
32	Non	A ouvrir	poles emploi,		Région	2		NPdC	bd à										
33	Non	A ouvrir	Implantation		Région	1	oui	NPdC	contacter la										
34	Non	A ouvrir	Autres		Région	1	oui	NPdC	exploitation										
35	Non	A ouvrir	Autres		Région	1	oui	NPdC	exploitation										
36	Non	A ouvrir	sites classés		Région	1	non	NPdC	extraction de										
37	Non	A ouvrir	zoos		Région	1	non	NPdC	extraction de										
38	Non	A ouvrir	Musées labellisés		Région	1	oui	NPdC	extraction de sirene										
39	Non	A ouvrir	Salles de		Région	1	oui	NPdC	extraction de										
40	Non	A ouvrir	Centres de		Région	1	non	NPdC	extraction de										
41	Non	A ouvrir	Télés Nord		Région	1	oui	NPdC	extraction de										
42	Non	A ouvrir	Forêts			3	non	NPdC	Intégration en										
43	Non	A ouvrir	nb de		Région	1	oui	NPdC	jointure avec										
44	Non	A ouvrir	BD Basias	Anciens sites industrielz et	Région / BRGM	3	non	NPdC	Récupérer la data et	BRGM	1/50 000	2012	B2						
45	Non	A ouvrir	Terrils		Région / EPF	2	non	NPdC	Récuprer donnée et à	*Chaîne des Textile*	*1/25 000*	*2008*	*D1*			*X*			
46	Non	A ouvrir	Carto		Région	3	oui	NPdC	réfléchir à la										
47	Non	A ouvrir	équipements			3	non	NPdC	valorisation										
48	Non	A ouvrir	Ligne de front		Région	3	oui	NPdC											
49	Non	A ouvrir	Lidar	Lidar Lille -	Région	2	non	Lille - Bassin									X	X	X
50	Non	A ouvrir	Régions	Périmètre	Région					IFN / ONF	1/200 000	2000	A3				X	X	X
51	Non	A ouvrir	Zones	Quartiers	Région						1/25 000	2006	A3				X	X	X
52	Non	A ouvrir	Zones	Quartiers	Région						1/25 000	2006	A3				X	X	X
53	non	A ouvrir	Fonds		Région					Préfecture de	1/50 000	2007	A3	R			X	X	X
54	Non	A ouvrir	Mosaïque	Résolution	Région					CNES® -	1/50 000	1990/91	A2	R			X	X	X
55	Non	A ouvrir	Photos	Rectification	Eurosense /						1 m	1997/98	C2	R			X	X	X
56	Oui	A ouvrir	résultats			2	oui	NPdC	bd à										
57	??	A ouvrir	Lien rapide			1	oui	NPdC	A intégrer										
58	??	A ouvrir	Canal Seine-			1	oui	NPdC	à intégrer										
59	A complétel	A ouvrir	bornes de			1	oui	NPdC	à constituer										
60	A complétel	A ouvrir	trafic			1	oui	NPdC	à constituer										
61	https://www	A ouvrir	Ecoles		Région	2	non	NPdC	of etalab										
62	https://www	A ouvrir	Politique de la	Périmètre des	Région					Préfecture de	1/50 000	**2007**	A3	R			X	X	X
63	https://www	A ouvrir	Grandes		Région	2	oui	NPdC	à constituer										
64	https://www	A ouvrir	établissement		Région	2	oui	NPdC	à constituer										
65	https://www	A ouvrir	équipements		Région	1	oui	NPdC	of bd										
66	https://www	A ouvrir	BD Histolitt	MNT marin	Région	2	non	NPdC	Intégrer								X	X	X

Données ayant été utilisées pour des cartes mais non disponibles sur le serveur potentiellement ouvrables

Annexe 5 : Présentation Open Data

Diapositive 1



Connaissance collaborative:

L'Open Data: l'ouverture des données (publiques)

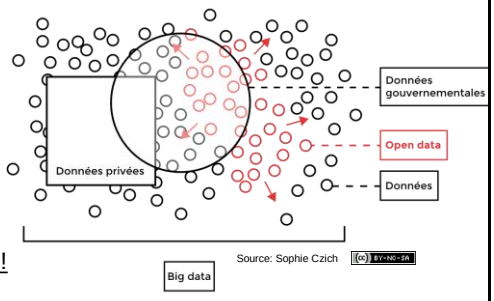
Louis PLATT
Stagiaire Open Data
Direction du Développement Durable, de la Prospective et de l'Evaluation
Louis.platt@nordpasdecalais.fr

Présentation dans le cadre de la rencontre 12/14
« la société du partage »
Mardi 2 juin 2015

Diapositive 2

L'Open Data c'est quoi?

- Démarche d'ouverture
- d'une donnée numérique
- d'origine publique ou privée
- diffusée de manière structurée
- selon une méthode et une licence ouverte
- garantissant son libre accès sa réutilisation par tous,
- sans restriction technique, juridique ou financière.



Source: Sophie Czich

Toutes les données ne sont pas ouvrables!

- Les données nominatives, à caractère personnel ou permettant de remonter à un individu
- Les données sensibles(sécurité nationale, secret défense)

Sont exclues de l'Open Data

Diapositive 3

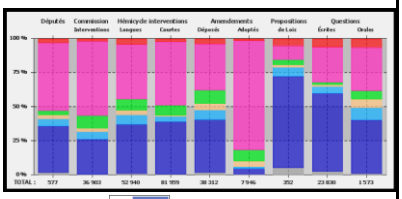
L'Open Data à quoi ça sert?(1/3):

Informer le citoyen et favoriser la transparence de l'action publique:

Source: wheredoesmymoneygo/

Répartition des taxes par domaine en fonction du salaire journalier au Royaume-Uni

Répartition de l'activité des députés sur les 12 derniers mois par groupe politique :



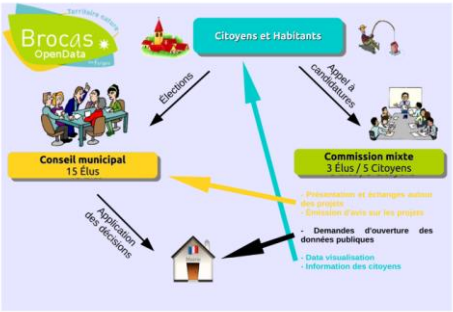
Source: Nosdéputés.fr/

Diapositive 4

L'Open Data à quoi ça sert?(2/3):

Moderniser l'action publique:

- décloisonner les administrations publiques;
- favoriser les échanges et la mutualisation entre les collectivités;
- favoriser la circulation des données au sein des administrations.;
- favoriser l'intégration des citoyens dans la construction des politiques publiques;



Source: Brocas 2012

Diapositive 5

L'Open Data à quoi ça sert?(3/3):

Favoriser l'innovation et le développement économique:

- Permettre le développement d'applications et de service innovants utilisant les données ouvertes.
- Favoriser la sérendipité(le hasard)

Source: handimap.org, ODbl

Diapositive 6

L'Open Data une démarche récente des autorités publiques…

Apparue il y a 6 ans en tant qu'enjeu politique
mais des initiatives antérieures dans la recherche ou le monde du libre

Les dates clés de l'Open Data :

internationales | nationales | Régionales

- 1995 : Apparition du terme Open Data aux USA
- 2009 : USA : Open Government Initiative
- 2010 : Royaume-Uni: lancement de data.gouv.uk
- 2011 : Mise en place de la mission Etalab;
- 2012 : Open Data comme principe gouvernemental
- 2013 délibération en faveur de l'ouverture, mise en place du site *http://opendata.nordpasdecalais.fr*

- avant 2006 : Sigale a partagé les données SIG sous convention très permissive
- 2010 : Rennes pionnier de l'Open Data
- 2012 : inscription au Comité Régional Numérique

Diapositive 7

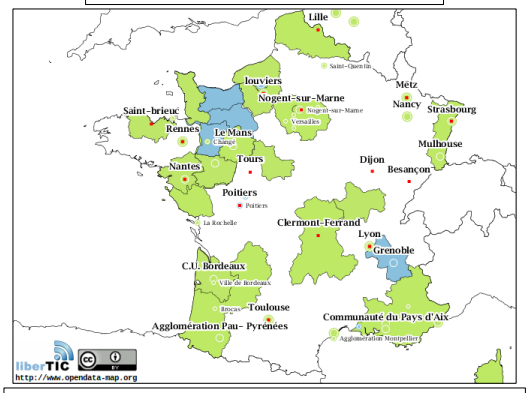
…qui trouve un écho particulier en France:

France: 3ème au classement mondiale de l'Open Data (Source: Open Knowledge Foundation)

Aujourd'hui:

- Plus de 50 projets Open Data issus des Collectivités Territoriales.
- De nombreux mouvements citoyens
- Une plateforme nationale comptabilisant:
  - 14 386 jeux de données ouverts
  - 531 organisations contributrices (dont 217 issues du service public)
  - 8963 réutilisateurs enregistrés
  - 1170 réutilisations des données référencées

Carte des initiatives Open Data en France:



Légende: Déjà ouvert En cours Mouvement citoyen

Diapositive 8

Le projet Open Data de la Région Nord-Pas De Calais(1/7):

Des choix politiques importants :

- La donnée publique comme **Bien Commun**
  - Intérêt général pour la donnée :
    - qui doit rester accessible à tous,
    - et sans frein
  - Renforce les valeurs de mutualisation, de coproduction
- L'Open Data comme **composante de la politique de Développement Durable :**
  - Faciliter la participation citoyenne
  - Transparence et co-construction de l'action publique
- L'Open Data comme **composante de l'action économique** :
  - Réemploi et valorisation économique
  - Favorise l'émergence de services innovants et utiles

Diapositive 9

Le projet Open Data de la Région Nord-Pas De Calais(2/7):

L'Open Data est intégrée à la politique numérique Régionale:

- Inscription dans le Schéma Directeur des Usages et Services du Numérique (SDUS)
  - Une Région en pointe sur le Numérique
  - Open Data = une opération en tant que telle
  - Mais aussi inscrite dans beaucoup d'autres fiches thématiques

Un pilotage institutionnel fort :

- Inscription au PDSI dans une opération majeure pour la visibilité institutionnelle
  - Enjeu de pilotage de l'action publique
  - Enjeu organisationnel : décloisonnement, transversalité
  - Opération 5 du PDSI « développement de la présence numérique de la Région »

Diapositive 10

Le projet Open Data de la Région Nord-Pas De Calais(3/7):

- Deux chantiers Open Data pour la Région:
  - Une démarche progressive axée en premier lieu sur l'interne:
    - Ne rien proposer aux autres qu'on ne sache mettre en œuvre soi-même
    - Mise en place d'un site web:
      https://opendata.nordpasdecalais.fr/
    - Rendre accessible les données de la Région
    - Développer l'animation et la communication autour de cette plateforme
  - Des actions incitatives auprès des acteurs intitutionnels du territoire à la libération de leurs propres jeux de données
    - Animation d'une communauté (Récolte) et des groupes de travail réguliers
    - Réalisation et diffusion d'une veille sur le sujet

Diapositive 11

Le projet Open Data de la Région Nord-Pas De Calais(4/7):

Une plateforme pour rendre les données accessibles à tous:

- Implication progressive des directions
  - expérimentation par 6 Directions en 2013
  - 50 jeux de données en 2013,
  - 30 nouveaux validés début 2015 (avec 4 nouvelles directions)
  - 82 jeux de données ouverts à ce jour
- Des exemples de réutilisations:
  - Des bureaux d'études réutilisent l'enquête ménage déplacement
  - Limites administratives (allers-retours : OpenStreetMap / Région)
  - Une direction récupère directement le budget sur le site Open Data

Diapositive 12

Le projet Open Data de la Région Nord-Pas De Calais(5/7):

Deux licences libres fixant les règles autour de l'utilisation de ces données:

La Licence Ouverte d'Etalab(LO), l'Open DataBase licence(ODbL): Elles **protègent le producteur de données, fixent les règles d'utilisation de la donnée.**

Vous êtes libres :

De partager : copier, distribuer et utiliser la base de données.

De créer : produire des créations à partir de cette base de données.

D'adapter : modifier, transformer et construire à partir de cette base de données.

En retour vous devez pour toutes utilisations publiques, ou pour toutes créations produites à partir de la base de données :

mentionner la source de la base de données, Indiquer la licence et garder intacte toute mention légale sur la base de données originaire.

spécifiquement pour la licence ODBL :

partager aux conditions identiques, les productions adaptées de la base de données

Garder ouvert : Obligation de proposer une version gratuite et libre des productions issues de la base de données

Diapositive 13

Le projet Open Data de la Région Nord-Pas De Calais(6/7):

Source: Etalab,

Une nécessaire démarche d'animation:

- Des événements annuels pour favoriser la réutilisation des données:
  - Hackaton
  - Cartoparties
  - Exemple: Etape du Ban TOUR (Base Adresse Nationale), Mars 2015)
- Communiquer lors des grands événements dont la Région est partenaires:
  - Mise en ligne de données spécifiques
  - Mises en place d'applications web cartographiques
- Communiquer pour sensibiliser et favoriser les réutilisations:
  - En interne (comme aujourd'hui)
  - Auprès des citoyens
  - Auprès des entreprises, et des acteurs économiques
  - Auprès des associations
  - Etc…

Diapositive 14

Le projet Open Data de la Région Nord-Pas De Calais(7/7):

Quels intérêts pour les agents de la collectivité?

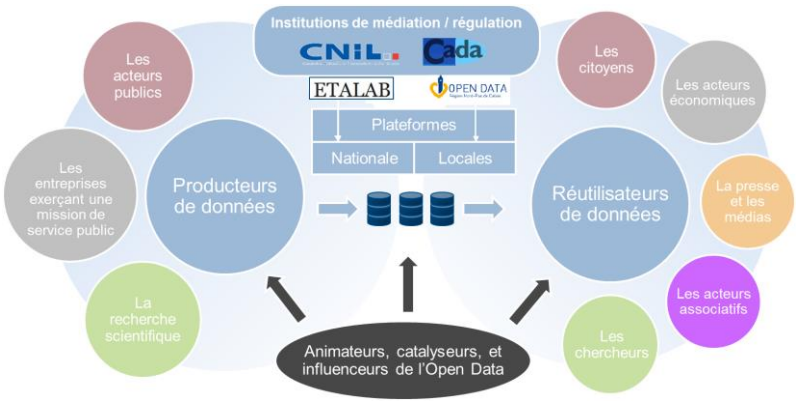
- Fluidifier l'information entre les services
- Disposer d'un jeu de données unique servant de référentiel commun
- Disposer rapidement des données à jour
- Moins de travail d'organisation et de gestion des données
- Une montée en qualité de la donnée
- Etc…

« Chers producteurs de données ouvertes, maintenant mangez-les ! »
Claire Gallon, Libertic

L'Open data ce n'est pas:

- **Une charge de travail supplémentaire**: à termes il s'agit de mettre en place des connecteurs entre la plateforme Open Data, les logiciels métiers des services et le système d'informations pour aller directement récupérer la donnée et la mettre à jour sur le site.
- **« Tendre le bâton pour se faire battre »:** il existe dans toutes les bases de données (publiques et privées)des erreurs, des imperfections. L'ouverture au plus grand nombre permettra de les repérer rapidement et aux données de monter en qualité.

Diapositive 15



Source: http://www.digitalcorner-solucom.fr droits de reproduction réservés et strictement limités.

Diapositive 16

Merci pour votre attention

Pour nous contacter:

Adrien Carpentier
Géomaticien,
adrien.carpentier@nordpasdecalais.fr
Tél. : 03.28.82.70.47
Service Information Géographique et Analyse Spatiale
Direction Développement Durable, Prospective, Evaluation

Louis Platt
Stagiaire Open Data,
Louis.platt@nordpasdecalais.fr
Tél. : 03.28.82.70.22
Service Information Géographique et Analyse Spatiale
Direction Développement Durable, Prospective, Evaluation

L'Open Data :

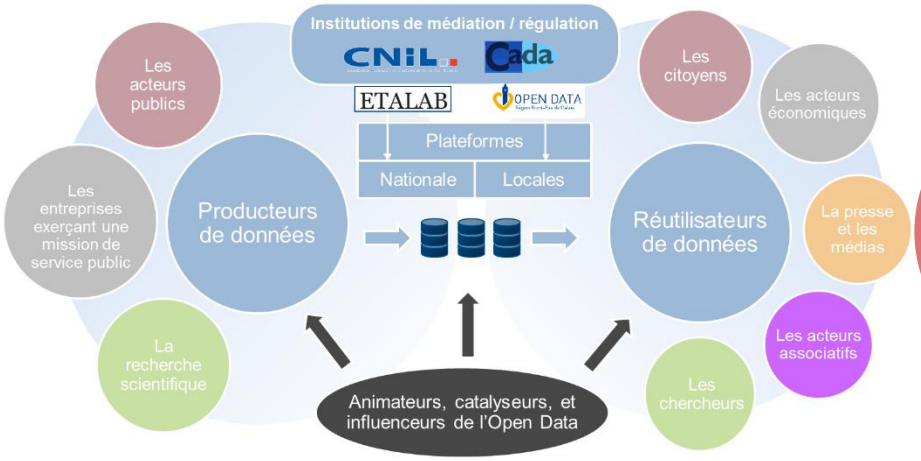
L'ouverture des données publiques, c'est…

Présentation Générale :

Définitions :

Démarche d'ouverture d'une donnée numérique,
d'origine publique ou privée, diffusée de manière structurée selon une méthode et une licence ouverte garantissant son libre accès sa réutilisation par tous, sans restriction technique, juridique ou financière.

Schéma explicatif de la démarche Open Data en France :



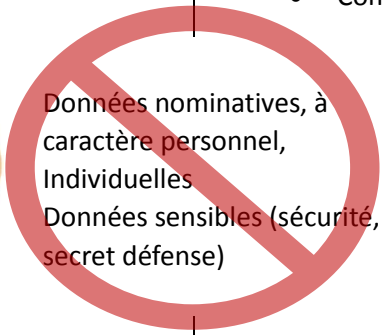
Source: http://www.digitalcorner-solucom.fr droits de reproduction réservés et strictement limités.

Les enjeux de l'Open Data :

- **Transparence de l'action publique**
- **Modernisation de l'action publique et participation citoyenne**
- **Innovation et développement économique**

Quelques chiffres clés :

6 ans que les premières initiatives publiques d'Open Data sont apparues en France.
Plus de **50** collectivités territoriales se sont lancées dans une démarche Open Data.
14 386 jeux de données ouverts sur la plateforme nationale
521 organisations contributrices sur la plateforme nationale



Données nominatives, à caractère personnel, Individuelles
Données sensibles (sécurité, secret défense)

A la Région Nord-Pas de Calais :

La donnée publique est un **Bien Commun**

Un enjeu politique et institutionnel fort :

- Inscription au Schéma de Développement des Usages et Services(SDUS)
- Inscription au Plan de Développement du Système d'information(PDSI)

Un outil au service des autres politiques :

- Transformation Ecologique et Sociale de La Région (TESR)
- Composante de l'Action Economique de la Région

Une plateforme Open Data accessible à tous :
https://opendata.nordpasdecalais.fr/

Quelques chiffres clés :

2 ans que la plateforme Open Data a été mise en place
10 Directions Impliquées
82 Jeux de données ouverts

Deux licences libres **qui protègent les producteurs** de données et **définissent le cadre d'utilisation** des données. La Licence Ouverte, l'Open DataBase Licence.

La mise en place progressive de dispositifs de communications et d'animations internes et externes pour sensibiliser à l'Open Data, faire connaitre, développer, favoriser les réutilisations de la plateforme et de ces données.

L'intérêt pour les Agents :

- **Fluidifier l'information entre les services**
- **Disposer de jeux de données uniques servant de référentiel commun**
- **Disposer rapidement des données à jour**
- **Moins de travail d'organisation et de gestion des données**

Vous souhaitez des précisions, ouvrir vos jeux de données, contactez-nous :

Adrien Carpentier
Géomaticien,
D2DPE/IGAS
adrien.carpentier@nordpasdecalais.fr

Louis PLATT
Stagiaire Open Data
D2DPE/IGAS
louis.platt@nordpasdecalais.fr

Annexes 7 : Note de synthèse donnée de santé

Ouverture des données de santé :

La France est l'un des pays disposant de la plus importante base de donnée de fichier santé au monde, le système d'information inter-régime de l'assurance maladie traite annuellement plus d'1,2milliard de feuilles de soins.

L'article 47 de la loi de modernisation du système de santé adoptée en première lecture le 14 avril par l'assemblée nationale prévoit l'ouverture des données publiques de santé. L'article 47 prévoit notamment la création du Système National des Données de Santé (SNDS), système regroupant un nombre très important de jeux de données de santé. Ce système de donnée sera géré par la Caisse nationale d'Assurance Maladie.

Quelles données :

La liste précise des jeux de données à ouvrir reste à définir. Cependant Les données que contiendra le système national des données de santé seront issues :

- des Systèmes d'Information des Etablissement de Santé, élaborées dans le cadre du Programme de Médicalisation des Systèmes d'Informations des Hôpitaux (PMSIH)
- Du Système national d'information inter-régime de l'Assurance maladie (Sniiram)
- Des données statistiques relatives aux causes de décès des Collectivités Territoriales
- Des systèmes d'information des maisons départementales des personnes handicapées (données « *médico-sociales* »)
- D'un « *échantillon représentatif des données de remboursement par bénéficiaire* » transmis par les mutuelles.

Comment ?

L'ensemble des données doit être anonymisé afin de ne permettre en aucun cas l'identification d'un individu avant l'ouverture au public.

La loi prévoit deux modes d'ouvertures différents

- Les données ne présentant aucun risque de ré-identification seront ouvertes à tous.
- Les données présentant un risque de ré-identification après recoupage pourront être mises à disposition de certaines associations, entreprises, journalistes après étude de leur méthodologie de traitement et de l'utilisation qu'ils souhaitent en faire par un comité d'experts scientifiques, un collège de l'Institut national des données de santé, la commission nationale de l'informatique et des libertés. Et enfin prouver que ces informations serviront des recherches d'intérêt public

Une attention particulière sera portée aux demandes provenant des sociétés d'assurance et des laboratoires pharmaceutiques, qui devront prouver que leur requête ne vise pas le ciblage de leur publicité.

Opportunités:

- Transparence de l'action publique.
- Amélioration des politiques publiques.
- Amélioration de la recherche en santé publique.
- Meilleur gestion humaine et financière dans le domaine de la santé.
- Prévention des risques liés aux médicaments.
- Regroupement avec d'autres données.

Limites:

- Nécessité accrue de protéger ces données dites « sensibles ».
- Potentielles dérives d'utilisation.
- Des données accessibles, mais pas ouvertes.
- Absence de définition de licence d'utilisation (Licence libre, ODBL …).
- Une ouverture limitée.

Conclusion :

L'article 47 de la loi de modernisation du système de santé national ouvre la voie à l'open data dans le domaine de la santé. Cependant cette ouverture semble trop « timide » du point de vue des défendeurs de l'open data, alors qu'elle soulève de nombreuses contestations et questions relatives au secret médical et au dérive lié au caractère personnel et sensible des données brutes.

L'ouverture des données de santé s'inscrit dans la volonté du gouvernement d'améliorer les politiques publiques et la transparence de l'action publique. Cependant cette loi, met en avant une ouverture partielle et très encadrée de ces données « sensible » et reste vague quant à ces modalités d'applications (quelle(s) licence(s)utiliser pour l'ouverture de ces données? quels jeux de données ouvrir ?). Néanmoins les expériences menées par ETALAB (hackaton + datascience donnée de santé) au début de l'année 2015 montrent bien l'intérêt des chercheurs et des professionnels et pour l'utilisation de ces données. On peut donc penser que ce projet de loi ouvre la voie à une ouverture plus large et mieux encadrée des données de santé dans les années à venir.

Bibliographie :

Commission Open data et santé, *Rapport final commission open data et santé*, rapport à Marisol Touraine ministre des Affaires sociales et de la Santé, 09/07/2014, www.drees.sante.gouv.fr/IMG/pdf/rapport_final_commission_open_data-2.pdf, 63p

Infographie, Ordre Nationale des Médecins, *données personnelles de santé,* novembre 2014http://www.conseil-national.medecin.fr/sites/default/files/cn_webzine/2014-11/www/index.php#/page-1(consulté le 15/04/15)

Les députés votent pour l'Open Data sur les données de santé, avril 2015http://www.nextinpact.com/news/93790-les-deputes-votent-pour-open-data-sur-donnees-sante.htm (consulté le 15/04/15)

Ouvrir les données de santé: la difficile équation du gouvernement, 09/04/15 http://www.lagazettedescommunes.com/347956/ouvrir-les-donnees-de-sante-la-difficile-equation-du-gouvernement/ (consulté le 15/04/15

PROJET DE LOI, *de modernisation de notre système de santé,* adopté en 1ère lecture le 14/04/15, http://www.assemblee-nationale.fr/14/ta/ta0505.asp

Que prévoit le projet de loi de santé concernant nos données médicales ?, 09/04/15http://www.francetvinfo.fr/sante/loi-sante/que-prevoit-le-projet-de-loi-de-sante-concernant-nos-donnees-medicales_870297.html#xtor=AL-79-[article]-[connexe] (consulté le 15/04/15)

Retour sur le premier Hackathon « données de santé », 02/02/15, https://www.etalab.gouv.fr/retour-sur-le-premier-hackathon-donnees-de-sante (consulté le 15/04/15)

Table des figures :

Figure 1 : Évolution de l'intérêt de l'Open Data selon les phases de développement……..p15

Figure 2 : liste des critères auxquelles doivent répondre les smart Cities (Giffinger 2007)…….....p19

Figure 3 : Processus classique de création de connaissance à partir de données brutes ……p20

Figure 4 : Processus de création de connaissance à partir des données métiers : ……p22

Figure 5 : Grille d'analyse des dispositifs d'animation de l'Open Data Open Data : …….p26

Figure 6 : Dispositif DataLab …….p28

Figure 7 : Le concept d'InfoLab ……p32

Figure 8 : Carte des dispositifs Open Data en France ……p38

Figure 9 : capture d'écran de la plateforme Open Data de la Région Nord-Pas de Calais ……p41

Figure 10 : Extrait du tableau de suivi des jeux de données ouverts (vert) et en attente de validation (jaune). ……p43

Figure 11 : Schéma récapitulatif des démarches Open Data en France……p50

Table des sigles :

AFIGéo : Association Française pour l'information géographique

AMO : Assistance à Maitrise d'Ouvrage

API ; Interface De Programmation

APPP : pôle Appui au Pilotage des Politiques Publiques(APPP).

CADA : Commission d'Accès aux Documents Administratifs

CNRS : Centre National de Recherche Scientifique

CNIG : Conseil National de L'Information Géographique

D2DPE : Direction du Développement Durable de la Prospective et de l'Evaluation

DAJC : Direction des affaires juridique et contentieux

DGA : Directeur Général Adjoint

DGS Directeur Général des Services

eGMS : e-Governement Metadata Standard

IGAS : service Information Géographique et Analyse Spatiale

IGN : l'Institut nationale de l'information géographique forestière

INRA : l'Institut National de la Recherche Agronomique

INSEE : Institue Nationale de la Statistique et des Études Économiques

LO : Licence Ouverte

NOTRe : Nouvelle Organisation Territoriale de la République

OAI: Open Acces Initiative

ODbl: Open Database License

OGL: Open Government Licence

OSM: OpenStreetMap

PPIGE : Plateforme Publique de L'Information GEographique

PSI : Public Sector Information

RECOLTE : REseau CoLlaboratif des acteurs et projets Territoriaux

SDUS : Schéma Directeur des Usages et Services numériques

SIG : Système d'Information Géographique

SIGALE : Système d'Information Géographique et de l'Analyse de l'Environnement

SRADDT : Schéma Régional d'Aménagement et de Développement Durable du Territoire

TER : Train Express Régional

TESR : Transformation Ecologique et Sociale de la Région

TIC ; Technologies de l'Information et de la Communication

TRI : Troisième Révolution Industrielle

Bibliographie :

Ouvrages scientifiques :

Aigrain (P.), 2005, Cause commune : l'information entre bien commun et propriété, éd. Fayard, coll transversale, 286p disponible sur : http://grit-transversales.org/IMG/pdf/Causecommune-CC-By-NC-ND.pdf

Anderson (C.), 2009, Free ! Entrez dans l'économie du gratuit, éd Pearson

Chebroux (J.-P.), 2007, Créer et animer un observatoire local, Territorial éditions, Voiron, 111 p.

Chignard S., 2012, *L'open data, comprendre l'ouverture des données publiques,* éd. FYP éditions, coll. Entreprendre, 192p

Roux (E.), Feyt (G.), 2011, Les Observatoires territoriaux : sens et enjeux, Ed. La documentation française, Coll. Travaux (DATAR), 112p.

Articles scientifiques :

Bacqué (M-H).), Sintomer (Y.), 2011, « Le temps long de la participation », in la démocratie participative, histoire et généalogie, éd.La découverte, coll. Recherches p1-36

Bacqué (M-H), 2005, « L'intraduisible notion d'empowerment vue au fil des politiques urbaines américaines », Territoires, n° 460, p. 32-35.

Bourcier (D.), De Filippi (P.), « L'Open Data : universalité du principe et diversité des expériences ? », in La Semaine Juridique Administrations et Collectivités territoriales, n° 13-2013, pp.1-9

De Sède-Marceau (M-H.), Moine (A.), Thiam (S.), 2009, Développement d'observatoires territoriaux, entre complexité et pragmatisme publié dans Rencontres SIG la Lettre, disponible sur ; http://www.sig-la-lettre.com/pdf-secure/R2009-Obs-desede.pdf

Denis (J.), Goëta (S.), 2013, « La fabrique des données brutes. Le travail en coulisses de l'open data » Manuscrit auteur, publié dans journée d'étude "Penser l'écosystème des données. Les enjeux scientifiques et politiques des données numériques », Paris disponible sur https://halshs.archives-ouvertes.fr/halshs-00990771/document

Feyt (G.), 2006, Redistribution des pouvoirs, redistribution des cartes : la connaissance des territoires, enjeu inédit de l'action publique ?, communication lors de Journée d'étude « LES POLITIQUES PUBLIQUES A L'EPREUVE DE L'ACTION TERRITORIALE », Grenoble disponible sur : http://www.pacte-grenoble.fr/wp-content/uploads/pdf_FEYT-.AFSP_.doc.pdf

Feyt (G.), Lardon (S.), 2010, L'émergence de nouveaux référentiels pour dire le territoire : vers une complémentarité des points de vue ou un « combat des cartes ?, Manuscrit auteur, publié dans "Colloque OPDE 2010 INRA Montpellier " Aide à la décision et gouvernance ", Montpellier, 9p. Disponible sur : http://hal.archivesouvertes.fr/docs/00/69/47/83/PDF/Article_communication_OPDE_Feyt_Lardon.pdf

Feyt, (G.), 2011 : « Les visages et usages de l'information géographique dans le processus de décision territoriale », in Joerin (F.), Schuler(M.), Thévoz (L.), Walser (O.), Joost(S.), Debarbieux (B.)(Dir.) *Les SIG au service du développement territorial*, Presses Polytechniques et Universitaires Romandes. 318p

Genoud (P.), Moeckli (A.), 2010, « Les tiers lieux, espaces d'émergence et de créativité », in La Revue économique et sociale, Vol 68 n°2 juin 2010, pp.25-33

Goëta(S.), Mabi(C.), 2014, « L'open data peut-il (encore) servir les citoyens ? », in Mouvements « (Contre-) pouvoirs du numérique »n°79, p81-91

Masoni-Lacroix (C.), Gherardi (L.), Rasse (P.), Batazzi (C.), Cadel (P.), et al., « Interroger l'accès et la réutilisation de données libres », in Les Cahiers de la SFSIC, n°10-2014, pp.176-178

Peugeot(V.) « Ouverture des données dans les collectivités territoriales : ambitions, racines politiques et premiers effets »,2012

Rive (M.), Noët (F.), Carmes (M.), Millet (L.), Meyer (F.), Noguier (M.), Raymond (P.), Boustany (J.), 2012, « L'information, un défi pour les collectivités territoriales », in *Documentaliste-Sciences de l'Information, n°49-2012 p. 44-59*

Roux (E.), 2012, « L'observation des territoires en marche », in intercommunalités n°170 p5-6

Volckrick, E., 2007, « Intervenir en tiers aujourd'hui », *Négociations*, 2007/1 n° 7, pp. 75-88, [en ligne], http://www.cairn.info/revue-negociations-2007-1-page-75.htm (Consulté le 01/05/2015)) 7

Rapports :

CEREMA, *L'Open Data en collectivité à la lumière des données de mobilité,* éd. Connaissance, mars 2015, 86p

Comité sur l'ouverture des données de transports présidé par Francis Jutand, *Ouverture des données de transports*, rapport au ministère des transports, de la pêche et de la mer, mars 2015, consultable en ligne : www.developpement-durable.gouv.fr/IMG/pdf/vf_rapport_jutand.pdf, 147p

Décider ensemble, *Ouverture des données publiques et participation : Quels enjeux démocratiques ?,* étude dans le cadre du projet « Ouverture des données publiques et participation : quels enjeux démocratiques ?, novembre 2012, disponible sur http://www.deciderensemble.com/images/stories/etudes/open_data/Ouverture_donnees_participation_etude_final.pdf 106p

Etalab, vade-mecum sur l'ouverture et le partage des données publics, septembre 2013, disponible sur : http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/vademecum-ouverture.pdf 11p

Giffinger R. et al., *Smart cities Ranking of European medium-sized cities,* rapport de recherché, octobre 2007, disponible sur : http://www.smart-cities.eu/download/smart_cities_final_report.pdf 27p

Mission commune d'information sur l'accès aux documents administratifs et aux données publiques, *Refonder le droit à l'information publique à l'heure du numérique : un enjeu citoyen, une opportunité stratégique*, rapport au Sénat, juin 2014, consultable en ligne : www.senat.fr/rap/r13-589-1/r13-589-11.pdf, 215p

Open Society institut et al*., Budapest Open acces initiative,* acte du congrès international de Budapest, décembre 2001, disponible sur : http://www.budapestopenaccessinitiative.org/read 1p

Philippe LEMOINE, *la nouvelle grammaire du succès-la transformation numérique de l'économie française,* rapport au gouvernement, novembre 2014, consultable en ligne : www.economie.gouv.fr/files/files/PDF/rapport_TNEF.pdf, 328p

TROJETTE (M A.) et al*., Ouverture des données publiques: Les exceptions au principe de gratuité sont-elles toutes légitimes ?,* rapport au premier ministre, juillet2013, consultable en ligne : www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/20131105 rapporttrojetteannexes.pdf , 121p

Autres:

Conseil National de L'Information Géographique, 2010, Présentation de la directive INSPIRE, consultable sur http://cnig.gouv.fr/?page_id=8991

Deslis (J.)2013, *Visualisation de données & Open data,* in support de cours Institut des mines Paristech Consultable sur : http://fr.slideshare.net/atbackbook/visualisation-de-donnes-open-data-intro

Loi n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal In Journal officie juillet 1978, p2851, consultable sur : http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068643&dateTexte=20150531

Malamud Carl Public.Resource.Org, *Program Open Government Working Group Meeting in Sebastopol, CA,* email au participants du working group sur l'Open Government, 27/10/2007 consultable sur: https://public.resource.org/open_government_meeting.html

Obama (B.), 2009, *Transparency and Open Government,* Memorandum for the Heads of Executive Departments and Agencies consultable sur: https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/

Czich (S.), 2015, Open data: forger une nouvelle culture, une nécessité pour le movement consultable sur: http://thesis.sophieczich.com/

Littérature grise :

Article de blog « Le citoyen a t-il une place dans l'open data ? », LiberTIC, 2012, https://libertic.wordpress.com/2012/02/17/le-citoyen-a-t-il-une-place-dans-lopen-data/ (consulté le 23/05/2015)

Article de blog « Vers la fin du baby blues de l'open data ? », LiberTIC, 2013 disponible sur : https://libertic.wordpress.com/2013/09/24/vers-la-fin-du-baby-blues-de-lopen-data/ (consulté le 15/05/2015)

Article de blog de datapublica interview de Claire Gallon sur le lancement du datalab pays-de-la-Loire, data publica 2013, http://www.data-publica.com/content/2013/09/interview-de-claire-gallon-sur-le-lancement-du-data-lab-pays-de-la-loire/ (consulté le 27/05/2015)

Article de blog : « Pourquoi partager ses données (quand la loi ne vous y oblige pas) ? », Chignard (S.), 2012, http://donneesouvertes.info/2012/06/ (consulté le 28/05/2015)

Article de la revue locatis : « Opendata France : les collectivités renforcent leur coopération », association OpenDataFrance 2013, http://www.localtis.info/cs/ContentServer?pagename=Localtis/LOCActu/ArticleActualite&jid=1250266087160&cid=1250266060334 (consulté le 24/05/2015)

Définition de données ouvertes, wikipédia, 2015, http://fr.wikipedia.org/wiki/Donn%C3%A9es_ouvertes(consulté le 15/052015)

Définition de l'Open Data, LiberTIC, 2010, In Actions Open data http://actionsopendata.org/l-opendata/definition/ (consulté le 01/05/2015)

Définition du logiciel libre, Free Software Foundation/GNU, 2015, http://www.gnu.org/philosophy/free-sw.html consulté le 20/05/2015)

Vidéos :

Open Data France, Audition des représentants Open Data France [ajoutée le 13/03/2014], audition dans le cadre des travaux de commission Mission commune d'information sur l'accès aux documents administratifs et aux données publiques, in Sénat,[49'09''] disponible sur : http://videos.senat.fr/video/videos/2014/video22153.html

RENCKI(J.), VERDIER(H.), Audition de MM. Julien RENCKI et Henri VERDIER, [ajoutée le 10/04/2014], audition dans le cadre des travaux de commission Mission commune d'information sur l'accès aux documents administratifs et aux données publiques, in Sénat,[60'21''] disponible sur : http://videos.senat.fr/video/videos/2014/video22418.html